

## THE ASSESSMENT OF CRITICAL EVALUATION, LEADERSHIP AND REFLECTION SKILLS THROUGH PARTICIPATION IN ONLINE DISCUSSIONS

Dr Jacqui Taylor

Psychology Research Centre, Bournemouth University, Poole. BH12 5BB

JTAYLOR@BOURNEMOUTH.AC.UK

### **Abstract**

Increasingly, educators from all disciplines are using blogs, social networking sites, vles and wikis to encourage academic discourse between students. However, a common problem experienced by educators is how these important learning experiences can be assessed and because of this difficulty many are not assessed. For some time now, i have been using online discussions via the University VLE as a way to encourage student debate around key lecture topics (e.g. Taylor, 2002). The key learning outcomes which this assessed activity addresses, in addition to learning more about the topic, are to develop skills in reflective practice, critical evaluation and leadership. Section 1 will review the ways that face-to-face and online academic discourse between students have been assessed, highlighting some of the differences to consider when setting up online discussion activities, compared to face-to-face discussion. Section 2 will then provide a case study of the way I set up online discussions and the method I currently use to assess contributions. The final part of the paper (section 3) will consider the potential for using quantitative content analysis (QCA) and automated methods to assess online participation.

### **Keywords**

*Online discussion, Online skill assessment, Virtual groups, Quantitative content analysis, Automating assessment*

### **1. Introduction**

Online group discussions develop many important traditional and 21<sup>st</sup> century skills important for graduates in this fast changing world. However, existing models of assessment typically fail to measure these skills. This section will consider the similarities and differences between the learning experiences in face-to-face and online group discussions and consider how the setting-up of online discussions can affect the development of skills.

## **1.1 Assessing ‘traditional’ and online discourse between students**

Academic discourse between students helps them to develop an understanding of different views on a topic and helps to develop public speaking and debating skills; however the learning that takes place during such interactions is rarely assessed. Although, students may be assessed individually for their presentation skills or for their contribution towards a group project, their actual discourse is rarely assessed, due mainly to the difficulty in assessing communication in real time face-to-face environments. Even if such interactions are recorded via video or audio, the complexity of such interaction makes assessment exceptionally time-consuming and difficult.

With discourse in online contexts, a record is made and interactions take place serially (rather than multiple parties communicating at the same time), and therefore it is easier to assess individual contributions. However, there is very little guidance for educators for assessing these interactions. Vonderwall et al (2009) highlight the paucity of research and practical advice on how to assess online postings. They discuss assessment processes in asynchronous online discussions, but focus on the variety of aspects that ‘could’ be assessed such as self-regulation, learner autonomy, learning community and student writing skills. They conclude that, ‘asynchronous online discussions facilitate a multidimensional process of assessment ...further research is needed to understand what assessment strategies or criteria enhance assessment and learning’ (p309).

In an early paper, Newman et al (1994) compared critical thinking in face-to-face and computer-based seminars (where students participated in both), and found that more new ideas emerged in the face-to-face seminars but that more ideas expressed in the online seminars were rated as important, justified or linked together; indicating a difference in quality and quantity. Heckman et al (2002) compared four face-to-face and four online group discussions and found that the online discussions generated high levels of cognitive activity, which were equal to or were superior to those identified in the face-to-face discussions. However, they provide no guidance on how to conduct assessment.

## **1.2 Setting up online academic discussions**

I have previously identified key factors for academics to consider when setting up online discussions and which affect participation and learning (Taylor, 2002). Three factors will be considered here which specifically relate to assessment (task type, assessment strategy, and individual differences), however other factors (such as group composition and instructor involvement) should also be considered.

As with offline learning activities, task type can have a significant impact on the quality and quantity of student engagement. Kanuka et al. (2007) examined the influence of different types of communication activity on the quality of students' messages in an online discussion and found significant differences between the five activities: the nominal group technique; debate; invited expert; WebQuest and reflective deliberation. They found that the WebQuest and debate activities led to the highest quality of messages and highlight the similar qualities of these two activities; that they were well structured, provided clearly defined roles for the students and they provoked the students to explicitly confront others' opinions. A number of research articles (e.g. Gafni & Geri, 2010) have shown that assessing online contributions both increases participation as well as enhances the quality of academic discourse. Swan et al. (2008) found that providing students with the assessment criteria led to increases in their participation and fostered deeper learning. Regarding the impact of individual differences, although gender and personality have been shown to affect preferences for online discussion, they did not affect performance. However, an individual difference to consider in assessment that has received little investigation in online learning is the impact of having English as a second language. Similarly, students who have communication difficulties such as dyslexia will need to be assessed according to the relevant marking guidelines.

## **2. Case study: my assessment of online discussions**

Hazari (2004) identified two ways to mark contributions to online discussions: analytic marking which involves assigning marks to specified criteria, and holistic marking where marks are assigned to the whole unit of analysis without scoring individual criteria. I use the analytical method and the unit of analysis that I use for assessment is the message; thus each message is evaluated using the criteria below. In addition to learning more about the topic and encouraging extended research, the key learning outcomes which the online discussion

activity addresses are to develop skills in critical reflection, critical evaluation and leadership. The assessment consists of participation in three online discussions and coordination and leadership of one discussion. Each individual message is assessed by hand using similar criteria to that used for other forms of academic writing, specifically whether it is analytical and evaluative. In addition, marks are awarded for reflection and timeliness of research cited. Each message is graded for these criteria using a five-point scale, from 'basic attempt' to 'excellent'. When leading an online discussion, students: provide an introduction to the key points; respond to other group members' questions; motivate discussion, and send a conclusion. The leader's posts are graded for coordination, responding and motivation. Some brief discussion of each of these skills highlights relevant research and the defining features for assessment.

## **2.1 Reflection**

Students are increasingly required to reflect critically on their learning as part of their coursework, however, as Coutinho (2007) highlights, teaching and encouraging reflective practice is problematic in many ways. For example, agreement on what constitutes reflective practice is vague and the assessment-centred approach to learning in HE often focuses reflection on improving the reflective writing style rather than on learning about learning (metacognition). In addition, when reflections are assessed, the incentive is to demonstrate knowledge and hide ignorance or doubt which is counter to Dewey's (1939) original purpose of reflection in which learning is derived from analysing mistakes and solving problems. Seale and Cann (2000) explored the way learning technologies were used to facilitate reflective thinking in students. They illustrated how a small group of students engaged with the material through online discussion; students were able to make links with other learning experiences and to see things in different ways. For my online discussions, students are encouraged to research widely using both academic sources as well as sources from the media (as long as due consideration is given to their credibility). This allows important concepts from published papers to be illustrated using clips from YouTube and the BBC, and web links to articles in the mass media. When I assess reflection, marks are awarded for links made to personal experiences and examples from the wider context such as news media.

## **2.2 Critical evaluation and extended research**

One of the key reasons for using online discussion is that it promotes deep learning, however this is difficult to measure, while measurement of critical thinking has been shown to be more reliable and occurs naturally in group discussion. Mason (1991) proposed that the measurement of online transcripts should be based on the educational value that they exhibited and broke this down into a number of useful questions, for example, whether a message built on previous messages, whether the participant drew on their own experience, whether they referred to course material or material outside the course and whether they initiated new ideas for discussion. When I assess evaluation, marks are awarded for questioning and building on previous messages and research. The extent and timeliness of research and resources used are recognised in marking, e.g. articles published within the last two years and for research not already covered in lectures.

### **2.3 Leadership skills**

In their study on the importance of trust in virtual teamwork, Jarvenpaa & Leidner (1998) found that teams with high trust levels were more capable of dealing with uncertainty and complexity than those with low levels. Prior to the online discussions, students are provided with a handout on the characteristics, benefits and problems of a virtual team, along with some useful tips for effectively managing and creating trust within their online discussions. Regarding the style of the online discussion, students are advised of the need to find a balance between social and task-based communication. The problems of coordinating or leading online teams can be significant, and leaders are encouraged to respond quickly and to include informal comments so that the style of discussion is not a series of long monologues and assessment identifies attempts to encourage interactivity, e.g. including participant's names and responding directly to them. This has been shown to be one of the key indicators demonstrating an awareness of social presence and community. A leader's role is critical at the start and end of a virtual discussion (when the definition of the topic and time plan are identified), therefore I assess the introduction and summary messages separately. Extra marks are awarded also for motivating comments and students are encouraged to weave the findings of empirical research into responses and questions to others.

### **3. Potential for using Quantitative Content Analysis (QCA) and software to assist assessment**

Due to a change to assessment strategy, the online discussions will form the only assessment for this unit in the future (currently it forms 30%, with 70% coming from an exam). Therefore there is a need to provide more detailed feedback to students regarding the assessment of their contributions and I am considering the potential for using content analysis or automated techniques as an additional method of assessment. A literature review of this area highlights the methodologies and software that could be used, and the limited empirical research highlights important factors to consider if assessment is based on these methods.

### **3.1 Quantitative content analysis (QCA)**

Newman et al (1995) developed a content analysis method to measure critical thinking in online group discussions and provided textual indicators to identify critical and uncritical thinking using sets of paired indicator, for example: relevant / irrelevant; important / trivial; new ideas / repeating what has been said; putting down new ideas / welcoming new ideas. The system certainly has potential for use in the assessment of online discourse as the use of obvious opposites should be easy for an academic (not experienced in qualitative analysis) to identify messages that illustrate these extremes. Other papers have been published since this, but they are not aimed at academics. For example, drawing on 19 key studies published in the preceding decade, Rourke et al (2001) cover the potential uses and the methodological challenges of analysing online transcripts using QCA. This classic paper provides a comprehensive discussion of issues relating to criteria, research designs, units of analysis and ethical issues, however it is not easy for the academic to use and seems primarily aimed at educational technologists and researchers. Indeed in a later paper, Rourke & Anderson (2004) propose that QCA is still not systematic and objective enough to describe academic discourse and provides procedures for developing the validity of a QCA coding protocol that is theoretically valid and to establish its validity empirically. Many of the articles employing QCA in online environments are theoretically driven by the Community of Inquiry (CoI) framework. This framework was developed by Garrison et al (2001) and consists of three elements: cognitive presence, social presence and teaching presence. The first two elements can be used to further understand the potential use of QCA in assessing reflection, evaluation and leadership.

#### **3.1.1 Critical evaluation and extended research**

Garrison (2001) provided a detailed overview of ways to evaluate online transcripts for evidence of critical thinking based on his five stages: (i) problem identification; (ii) problem definition; (iii) problem exploration; (iv) problem evaluation / applicability, and (v) problem integration. While Garrison's stages are useful they need some simplification if they are to be used to assess online discussions. It is clear that the main focus of later work by Garrison et al (2006) is in producing a methodology to systematically and rigorously measure cognitive presence in online communications. Their work is very useful in guiding educators in the adoption, design and implementation of online environments for learning, but is less useful as an assessment tool. Kanuka et al. (2007) used the construct cognitive presence to investigate the role of critical discourse in distance education and examined the influence of different types of communication activity on the quality of students' messages in an online discussion. Using QCA to analyse messages from 19 students in an undergraduate course, each message was assigned to one of the four categories of cognitive presence. While the number of contributions categorised in the highest phases of cognitive presence was low (20.21%), interestingly they found that it was highest during activities which were well structured, provided clearly defined roles for the students and they provoked the students to explicitly confront others' opinions. Extended research is probably the easiest to assess using QCA and to some extent can even be partially automated. For example, dates can be highlighted to allow easy identification of recent research and using a list of references already used in lectures the extent of new research cited by students can also be easily identified.

### **3.1.2 Reflection**

A study by Mair & Taylor (2011) set out to identify whether students were reflecting and if so, how deeply. A content analysis was conducted on discussion transcripts using the four types of reflective writing identified by Hatton and Smith (1995). The study found that the level of reflections within the postings became deeper over time. For example, although more reflection was occurring in the early discussions, the majority of the postings were classed according to Hatton and Smith (1995) as level 1 (merely reports events with no attempt to provide reasons) and level 2 (provides reasons, often based on personal judgement). While later discussions contain deeper, dialogic and critical reflections, i.e. more at Hatton and Smith's level 3 (discourse with one's self, mulling over reasons, exploring alternatives). Although this study highlighted the method as a potential way to categorise reflections, it was not used for assessment, due to time constraints.

### **3.1.3 Leadership skills**

A key skill of an online leader is to encourage an atmosphere of trust and collaboration and this has been linked to the concept of social presence. Rourke et al (1999), drawing on the community of inquiry work above, have produced a template to assess social presence through content analysis. The usability of the template for educators is enhanced through the provision of selections of coded transcripts and inter-rater reliability figures illustrate the validity of this template. However, this article is of most use for conference moderators and researchers as the focus is on setting-up and encouraging social presence, rather than the assessment of this factor.

### **3.2 Software to assist analysis, assessment and feedback**

Over the last five years, there has been an explosion in the use of computer software to analyse text and there are literally hundreds of software products available which can assist analysis of online text (e.g. kdnuggets). Many packages are developed for use in specific fields and contain features appropriate to the type of discourse being analysed, for example: politics (e.g. Hopkins & King, 2010 look at political speeches and campaigns); health (e.g. Kim, 2009 evaluated cancer blogs), and those commercially available for marketing and advertising. There are relatively few packages developed for pedagogic use or for use in the social sciences. A review of papers published over the last two years found that the packages used most often in the social sciences include: Linguistic Inquiry and Word Count (LIWC); QSR NVivo; TAMS Analyzer (Text Analysis Mark-up System); ATLAS.ti, TextSTAT, Ranks NL. These range in features, from those which produce word frequency lists and concordances to those with powerful search possibilities (e.g. to identify regular expressions or phrases).

Despite the advances of methods for automated content analysis in the field of media analysis, most methods are only able to highlight and count instances of pre-specified words or phrases and we are a long way from automated assessment of critical thinking for example. One potential package that could be used immediately is the free, open-source template NodeXL, which makes it easy to develop network graphs from data entered within a Microsoft® Excel® spreadsheet. This package has great potential to provide assessment feedback in a visual format.



#### **4. Conclusion**

It is clear that assessment methods need to be modernised, to reflect the changes in learning activities taking place as a result of using interactive and collaborative technologies made possible by Web 2.0. Additionally, updated methods need to consider the experiences and expectations of the current generation of students (Taylor, 2011). McGaw (2009) highlights a number of new 21<sup>st</sup> Century skills which are developed through interactions using social media. However, a theme throughout this article is that the assessment of these skills remains a challenging area for academics. The potential for partially automating assessment through the use of QCA has been proposed, however we are a long way from this. In addition to the difficulties in operationalising concepts such as reflection and evaluation, new systems need to account for the different nature and style of online academic discourse, e.g. the different level of formality, different cultural styles of language and to treat fairly those students with additional learning needs (such as dyslexia).

#### **References**

- Coutinho, S. A. (2007). The relationship between goals, metacognition and academic success. *Educate*, 7(1), pp.39-47.
- Dewey, J. (1939). Experience, knowledge and value: a rejoinder. In *The Philosophy of John Dewey*, edited by P. Schilpp, Northwestern University, Evanston (IL), pp. 517-608.
- Gafni, R., & Geri, N. (2010). The value of collaborative e-learning: Compulsory versus optional online forum assignments. *Interdisciplinary Journal of E-Learning and Learning Objects*, 6, pp.335-343.
- Garrison, D.R., Anderson, T. & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15(1), pp.7-23.
- Garrison, D. R., Cleveland-Innes, M., Koole, M. & Kappelman, J. (2006). Revisiting methodological issues in the analysis of transcripts: Negotiated coding and reliability. *The Internet and Higher Education* 9(1), pp.1-8.

- Garrison, D. R., & Akyol, Z. (2011). Understanding cognitive presence in an online and blended community of inquiry: Assessing outcomes and processes for deep approaches to learning. *British Journal of Educational Technology*, 42(2), pp.233-250.
- Hatton, N. & Smith, D. (1995). Reflection in teacher education. towards definition and implementation. *Teaching and Teacher Education*, 11, pp.33–49.
- Hazari, S. (2004). Strategy for the assessment of online course discussion. *Journal of Information Systems Education*, pp.349-355.
- Heckman, R. & Annabi, H. (2005). A content analytic comparison of learning processes in online and Face-to-Face Case Study Discussions. *JCMC*, 10(2).
- Hopkins, DJ & King, G. (2010). A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science*, 54(1).229-247.
- Jarvenpaa, S., & Leidner, D. (1998). Communication and trust in global virtual teams. *Journal of Computer Mediated Communication*, 3(4).
- Kanuka, H., Rourke, L. and Laflamme, E. (2007). The influence of instructional methods on the quality of online discussion. *British Journal of Educational Technology*, 38, pp.260–271.
- Kim, S. (2009). Content analysis of cancer blog posts. *Journal Medical Library Association*, 97(4), pp.260–266.
- Mair, C. & Taylor, J., (2011). Critical reflection and discussion facilitated by a virtual learning environment across two universities. *Paper presented at the International Technology, Education and Development (INTED) Conference, 7-9 March, Valencia, Spain.*
- Mason, R. (1991). Evaluation Methodologies for Computer Conferencing Applications. In A. R. Kaye, (Ed), *Collaborative Learning through Computer Conferencing*. Springer-Verlag, Berlin.
- McGaw, B. (2009). Transforming Education: Assessing and Teaching 21st Century Skills. [http://education2020.wikispaces.com/file/view/Transformative\\_Assessment--A\\_Call\\_to\\_Action\\_and\\_Action.pdf](http://education2020.wikispaces.com/file/view/Transformative_Assessment--A_Call_to_Action_and_Action.pdf) (accessed 9/2/12)
- Newman, G., Webb, B., & Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology*, 3(2), pp.6-77.
- 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 in Education*, 12, pp.8-22.

- Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, 52 (1), pp.5-18.
- Rourke, L., Anderson, T., Archer, W., & Garrison, R. (1999). Assessing social presence in asynchronous computer conferencing transcripts. *Journal of Distance Education*, 14(2), pp.50-71.
- Seale, J. K., & Cann, A. J. (2000). Reflection on-line or off-line: the role of learning technologies in encouraging students to reflect. *Computers and Education*, 34, pp.309-320.
- Swan, K., Shea, P., Richardson, J., Ice, P., Garrison, D. R., Cleveland-Innes, M., & Arbaugh, J. B. (2008). Validating a measurement tool of presence in online communities of inquiry. *E-Mentor*, 2(24), pp.1-12.
- Taylor, J. (2002). A review of the use of asynchronous e-seminars in undergraduate education. In *The Digital University: Building a learning community*, edited by R. Hazemi and S. Hailes. Springer, London. pp.125-138.
- Taylor, J. (2011). How social media is changing students and the way they learn ...and what psychology academics can do about it! *Paper presented at the International Conference on the Teaching of Psychology (ICoTP)*, July 29-31, Vancouver, Canada.
- Vonderwell, S, Xin Liang, X & Alderman, K. (2009). Asynchronous discussions and assessment in online learning. *Journal of Research on Technology in Education*, 3(3), pp.309-328.

#### Online resources:

- Kdnuggets: Text Analysis, Text Mining, and Information Retrieval Software.  
<http://www.kdnuggets.com/software/text.html> (accessed 6/1/12)
- Linguistic Inquiry and Word Count (LIWC) <http://www.liwc.net/> (accessed 6/1/12)
- TextSTAT <http://neon.niederlandistik.fu-berlin.de/en/textstat/> (accessed 6/1/12)
- QSR NVivo <http://www.qsrinternational.com> (accessed 6/1/12)
- ATLAS.ti <http://www.atlasti.com/> (accessed 6/1/12)
- Ranks NL: <http://www.ranks.nl/> (accessed 6/1/12)
- TAMS (Text Analysis Mark-up System) <http://tamsys.sourceforge.net/> (accessed 6/1/12)