Analysing online discussions: What are students learning?

Deborah Cotton, Jon Yorke
Educational Development and Learning Technologies
University of Plymouth

Online asynchronous discussions (OADs) are increasingly advocated to encourage interaction in blended learning in higher education. However, questions remain over the educational utility of OADs. In particular, relatively little is known about how students use online discussions and the ways and extent to which their use enhances learning. Previous research seeking to investigate the correlation between discussion board use and exam results has proved problematic and open to misinterpretation. Analysis of the content of online discussions may provide a more fruitful way of discovering the impact on student learning, but this approach can appear overly complex and time-consuming. This paper describes a small scale research project which pilots a number of different methods for analysing online discussions and considers the advantages and disadvantages with each approach, both in terms of methodological simplicity and utility of findings.

Keywords: web-based education, online discussion, blended learning, student learning, social presence

Introduction

Online asynchronous discussions (OADs) are increasingly used in higher education, often as part of a blended learning approach combining computer mediated and face to face interaction. Computer mediated discussions have the potential to provide opportunities for interaction and collaboration between learners, and to encourage informal peer or tutor-led learning opportunities at a number of levels. This offers the advantage of enabling interaction at a time and place convenient to the learner, and of supporting reflection on face-to-face sessions. However, the ways in which students engage in online discussions are likely to influence the learning outcomes achieved, and research which focuses on actual use of OADs (as opposed to an ideal view of what they might be used for) is therefore crucial.

It has been argued that online learning encourages wider student participation and increases interaction between students when compared with traditional programmes. However, the evidence of impact on student learning is far from clear-cut. Davies and Graff (2005) examined the frequency of online interactions of a group of undergraduate students and compared this with their end of year grades. Their findings suggested that greater online interaction did not necessarily translate into higher grades – although they did find that students who failed in one or more modules had interacted less frequently than those who passed. However, interpretation of these findings is problematic: Is it simply that the more able or strongly motivated students contributed more substantively to online discussions? Was it the case that those with a limited understanding (who subsequently failed) interacted less because they were already struggling with the subject area? These kinds of correlations, we believe, cannot give definitive answers to the question of the impact of online engagement on student learning. We suggest that a more fruitful approach to assessing the educational utility of OADs is to investigate and analyse the content of online discussions with respect to student learning.

Methodology

A number of authors have developed theoretical models of student learning through online discussion, and it is these theories which guide the current study. Previous research on online discussions has drawn upon a variety of methods, from simple counting of frequency of contributions (Davies & Graff, 2005); analysis of student perceptions of social presence (Richardson & Swan, 2003); or individual or team categorisation of statements within postings (Gilbert & Dabbagh, 2005; Murphy, 2004).

In this research, we piloted three different methods of analysing online discussions, in order to investigate the following questions:

- How easy is each method to implement?
- Do any methodological problems arise?
- To what extent does each method provide reliable and valid data about student learning?

The OAD used for the purposes of analysis was a recent online conference used to support the General Teaching Associates (GTA) course at the University of Plymouth in the United Kingdom. The GTA programme aims to support new and part time teaching staff (such as graduate students with limited teaching responsibilities). Successful completion of this 20 credit level 3 module leads to Registered Associate Practitioner status with the UK Higher Education Academy. Participants in this study were 17 GTA students and five tutors, and online discussions were based around specific tasks and activities. This particular version of the GTA course involves online activities scheduled between the taught sessions, and involvement in the online discussions is a required part of the course. Discussion comments posted by individuals receive formative assessment from peers and tutors, but are not subject to summative assessment.

Participants and tutors on the GTA course were invited to be involved in the research at the start of the programme, and their consent was obtained for use of the discussion transcripts. Ethical approval was obtained following standard university procedures, and an ethics protocol was developed.

Analysis of OADs

A sample of conversations from the March 2006 GTA conference were selected for use in the analysis. The aim was to utilise a variety of discussions involving both students and tutors, based around three specific activities and the generic learning log. The activities included discussions around 'learning styles', 'difficult situations' and 'assessment'. Within each of these strands and the learning log we selected two conversation threads to analyse. The criteria for selection were that conversations should contain postings by two or more participants, at least one of whom must be a tutor. In total therefore, eight conversation threads were analysed. More detail about each conversation is given in Table 1.

Table 1: Summary of participants and postings in selected conversation threads

Discussion Thread	Student participants	Tutor participants	Total participants	Student postings	Tutor postings	Total postings
Learning Styles 1	2	1	3	3	1	4
Learning Styles 2	3	1	4	5	1	6
Difficult Situations 1	3	2	5	3	2	5
Difficult Situations 2	4	2	6	8	2	10
Assessment 1	2	1	3	4	1	5
Assessment 2	2	1	3	4	2	6
Learning Log 1	5	1	6	6	1	7
Learning Log 2	3	1	4	4	1	5

The three methods of analysis piloted in this study were (i) analysing social, teaching and cognitive 'presence'; (ii) content analysis based on Bloom's taxonomy; and (iii) Quantitative analysis based on intended learning outcomes.

Method 1: Analysing social, teaching and cognitive 'presence'

Garrison, Anderson and Archer (2000) describe a model depicting three key dimensions of the learners' educational experience when using text based computer conferencing. These are *cognitive presence* (the extent to which participants can "construct meaning through sustained communication"), *teaching presence* (the design of the learning experience and facilitation both by tutors and students during a discussion) and *social presence* (the ability of participants "to project their personal characteristics into the community"). They assert that when teaching and social presence are both high, there is a positive impact on cognitive presence leading to effective learning and enhanced academic performance. This

claim is supported to some extent by Volet and Wosnitza (2004) who maintain that a strong sense of social presence contributed to the level of engagement amongst participants in their study of online discussions. However, Murphy (2004) explored the potential for student collaboration through online discussion, and found that participants engaged mainly in processes related to social presence and individual perspectives, concluding that more explicit scaffolding was required in order to encourage a stronger cognitive dimension. Our research involved an investigation of different aspects of 'presence' via an analysis of textual units drawn from an online discussion of students on the GTA programme. Based largely on the work of Garrison et al. (2000), we focused on three broad themes of social, teaching and cognitive presence, using the indicators outlined in Table 2.

Table 2: Indicators of presence (adapted from Garrison et al., 2000)

Aspect	Indicators
Teaching Presence	Selection, organisation, and primary presentation of course content The design and development of learning activities and assessment Facilitation (teacher and student)
Social Presence	Emotional expression Projection of personal characteristics
Cognitive Presence	Sharing of knowledge and ideas Negotiation of conflicting views

Using these indicators as broad themes, we categorised textual units posted in the online discussion, with each transcript being coded independently by two researchers in order to compare findings and gauge levels of inter-rater reliability. We calculated the quantity of each aspect in different conversation threads to investigate the relative densities of teaching, social and cognitive presence.

Method 2: Content analysis based on Bloom's taxonomy

This approach follows the methodology utilised by Gilbert and Dabbagh (2005) in their analysis of online discussions. These researchers created a coding scheme which assessed whether students were:

1 Relating new knowledge to prior knowledge
2 interpreting content through the analysis, synthesis and evaluation of others' understanding
3 making inferences

Codes used in this research are listed in Table 3 together with the mapping to Bloom's taxonomy (Bloom, 1956) suggested by the authors. This coding scheme was also piloted independently by two researchers in order that inter-rater reliability could be assessed.

Table 3: Coding scheme based on Bloom's taxonomy (adapted from Gilbert & Dabbagh, 2005)

Code name	Brief definition	Bloom's taxonomy reference
Reading Citation	Citation of set reading, e.g. reference to article or chapter by learner	Knowledge
Content clarification	Personal interpretation of content, e.g. paraphrasing concept or principles	Comprehension
Prior knowledge	Use of prior knowledge and outside resources to support statement	Comprehension
Real world example	Citing personal experience (professional/ academic) to demonstrate application to real-world context	Application
Abstract example	Use of analogies, metaphors or philosophical interpretations to support understanding	Application
Making inferences	Going beyond information given: beyond comprehension - adding or constructing new knowledge	Analysis, synthesis and evaluation
Facilitator question	Question posted by facilitator	n/a
Facilitator response	Response posted by facilitator	n/a
Facilitator clarification	Clarification posted by facilitator	n/a
Instructor posting	Messages posted by the instructor	n/a

Method 3: Quantitative analysis based on intended learning outcomes

To provide a simple comparison with the fairly complex coding schemes outlined above, we utilised a straightforward quantitative word-count method to assess the quantity of on-task discussion (i.e. that which directly related to the learning outcomes for that task), using the same OADs as above. This method involved simply counting the number of words in each discussion which were deemed to address the intended learning outcomes (as judged by two independent researchers). This was intended to provide a basic measure of the quantity of on-task discussion to compare with the two previous coding schemes which looked at differing aspects of quality of discussion.

An example of coded text illustrating all three methods is provided below. The learning outcomes for this activity asked learners to:

- Describe the VARK questionnaire and their results
- analyse their results in the light of their approaches to learning
- identify implications for their approaches to teaching.

Table 4: Example coding of student posting using 3 methods (Researcher 1)

Conversation Text	Coding Method 1	Coding Method 2	Coding Method 3
Hi	Organisational	Organisational	Off-task
Do you mean that you like to be very interactive with the students? Sharing love and joy in learning.	Cognitive presence	Off task	Off-task
I relate to this and I see the lab as a good place for student interaction.	Cognitive presence	Content clarification (personal interpretation)	Off-task
I have had good feedback from the students, not based on my knowledge of subject but on how I relate to them.	Cognitive presence	Real world example (personal experience)	Off-task
For me to become a lecturer would be interesting, I would have to bring the same interaction to the classroom.	Cognitive presence	Content clarification (interpretation of content)	On-task
The lab is definitely my environment, yet the lecturers give the labs to me because they hate labs.	Cognitive presence	Real world example (personal experience)	On-task
It is also the best place to judge the students abilities, exams are just about memory. Labs require students to demonstrate their skills and is similar to their future workplace.	Cognitive presence	Prior knowledge	Off-task
Sorry I have drifted off-topic a bit.	Social presence	Off-task	Off-task
Regards	Organisational	Organisational	Off-task

Results

The focus of this section is to illustrate the type of data produced via the three different methods (methodological issues are addressed later in this paper).

To recap, three questions guided this study with respect to the analysis of online discussions:

- How easy is each method to implement?
- Do any methodological problems arise?

- To what extent does each method provide reliable and valid data about student learning?

Method 1: Analysing social, teaching and cognitive 'presence'

This method of analysis produced a reasonable level of inter-rater reliability, with both researchers classifying the majority of the conversation as 'cognitive presence', some as 'social presence' and very little as 'teaching presence' (Table 5).

Table 5: Percentage of conversation in each category

Conversation	Teaching R1 %	Teaching R2 %	Cognitive R1 %	Cognitive R2 %	Social R1 %	Social R2 %
Assessment 1	6.3	6.3	78	64.3	15.5	29.2
Assessment 2	9.0	4.1	84.1	81.0	6.8	14.7
Difficult situations 1	0.0	1.9	72.1	60.5	27.8	37.5
Difficult situations 2	0.0	0.0	54.6	61.1	45.3	38.8
Learning logs 1	0.0	0.0	53.1	56.5	46.8	43.4
Learning logs 2	0.0	0.0	72.4	58.8	27.5	41.1
Learning styles 1	6.3	0.0	79.3	86.9	14.3	13.0
Learning styles 2	12.4	6.2	69.2	72.9	18.2	20.7

Note. R1 = 1st researcher; R2 = 2nd researcher

An interesting finding from this method was that social presence was rated as being higher in those activities with less teaching presence ('learning logs' and 'difficult situations'), and those which depended more on student experiences than teacher authority or prior learning. These two types of conversation were more reflective and personal than the 'assessment' and 'learning styles' tasks. However, it is not clear that much useful information about student learning is gained through this approach. The high volume of cognitive presence suggests that all of these OADs were enhancing student learning, but the extent of high-level cognitive activity cannot be gauged from these results.

Method 2: Content analysis based on Bloom's taxonomy

This method appears to have the potential to provide greater insights into student learning since it offers the opportunity to evaluate the level of cognitive engagement of the students involved in the discussion. Bloom's taxonomy can be used to categorise the level of understanding from the lower levels of knowledge and comprehension to the higher levels of analysis, synthesis and evaluation.

Table 6: Percentage of conversation in each category

	1	2	3	4	5	6	7	8	9	10	11	12
	R1 % read	R2 % read	R1 % cont	R2 % cont	R1 % prior	R2 % prior	R1 % real	R2 % real	R1 % abstr	R2 % abstr	R1 % infer	R2 % infer
Assess 1	5.9	0	18.4	6.4	0	0	8.5	40.5	0	10.1	29.7	18.6
Assess 2	0	0	34.3	10.7	0	0	4.1	0	0	6.2	11.8	35
Diff 1	0	0	0	0	0	3.2	53	49.8	0	8	11.2	0
Diff 2	0	0	0	3.2	0	0	63.3	40.1	0	25.3	21.4	17.6
L Logs 1	0	0	21.4	29.5	0	0	36	26.4	0	15.1	34.1	21.8
L Logs 2	0	0	6.3	16.5	0	8.8	21	11.3	0	6.9	19.8	5.8
L Style 1	0	6.2	27.4	11.8	9.7	5.9	21	34.2	4.2	3.6	10.2	18.1
L Style 2	0	3.8	14.9	23.9	8.5	0	23.7	11.9	0	23.4	21.3	12.5

There were some interesting findings from this approach to analysis. For example, both researchers noted the relatively low level of use of 'prior knowledge' (columns 5 and 6) and 'reading citation' (columns 1 and 2) in all conversations, and a relatively high level of 'real world examples' (columns 7 and 8). A relatively high proportion of the discussion was rated as 'making inferences' – the highest order category

in this scheme (encompassing analysis, synthesis and evaluation). However, as noted below, there were also significant differences between the categorisations of the two researchers.

Method 3: Quantitative analysis based on intended learning outcomes

This method aimed to isolate not solely cognitive activity, but the proportion of cognitive activity which was focused on the intended learning outcomes for each online activity. Unlike the previous method, no attempt was made to judge the level of student engagement. It is difficult to comment on the findings since the inter-rater reliability was very low. It appears that the learning outcomes were interpreted in a far more stringent manner by the module leader (R1), than by the other researcher (R2).

Table 7: Percentage of conversation related to learning outcomes

	% related to learning outcomes R1	% related to learning outcomes R2	Rank order R1	Rank order R2
Assessment 1	37.3	66.7	5	3
Assessment 2	14.3	50.3	8	7
Difficult Situations 1	51.3	57.8	1	5
Difficult Situations 2	39.5	57.1	4	6
Learning Logs 1	35.2	82.2	6	1
Learning Logs 2	33.4	49.3	7	8
Learning Styles 1	50.9	71.3	2	2
Learning Styles 2	47.7	61.2	3	4

Methodological issues

This study has raised a number of issues in relation to each of these approaches, both generic and specific in nature. Firstly, both researchers felt that the time taken to process these (relatively short) conversations was significant. Method 3 was the quickest to implement (approximately one hour), but clearly this approach also produced significant disagreement. Methods 1 and 2 were more time consuming, taking approximately two to three hours each. Method 2 was judged to be the most involved rating procedure to apply, though in practice the time taken by both researchers for this was similar to that of Method 1. However, this may have been due to prior familiarity with the material, something that would not necessarily be the case if Method 2 was adopted as a sole measure.

Other related issues that arose included the time taken to recognise and deal with incorrect posting of related content in a wrong discussion folder or conversation thread. Aspects that further hinder the application of methods such as these include participants' failure to include a subject line, duplication of subject titles and the varying ability of conferencing software to process and present conversation threads in a coherent manner. It was recognised that private email between individual students may also be present, and this may be making an invisible contribution to the learning outcomes. All of these issues provide a degree of impact on the time taken and the reliability of subsequent analysis.

It became clear that there is a need to measure 'off-topic' and 'organisational' aspects of conversation threads separately in discussion forums. Conversations usually have message headers, signatures and other extraneous information and these 'organisational' aspects should be discounted from measures of word count. They should not, however, be recorded as 'off-topic' conversation since this would suggest a particularly high level of irrelevant talk. Whilst including these aspects in the total word count makes the calculation easier, this produces a 'signal to noise' ratio (in the sense of on and off-topic conversation) that we feel is misleading. Specific issues related to each method are detailed below.

Method 1

Whilst Method 1 appears to produce reasonable inter-rater agreement, a number of issues arose in relation to the overlap between categories. The categories of social and cognitive presence were judged to be

particularly problematic, as a number of postings projected a social presence through comments that were judged to have a strong cognitive element. In this sense, social and cognitive presence may form a continuum, with classification being especially difficult towards the central point.
The concept of 'teaching presence' proved to be similarly problematic, in that a number of postings by tutors were judged to be primarily cognitive in terms of this classification structure. This was exemplified by one particular posting that provided an example of very subtle task direction by a tutor who engaged participants in relatively high level discussion. Had these postings been blind reviewed it is likely that this would have been classified as cognitive presence, teaching presence would not have been considered. It is likely that bias is present here, as there is a natural tendency to look for teaching presence where a discussion thread is known to be posted by a tutor.

It is also acknowledged that certain activities are specifically designed to encourage social presence and the formation of a community of practice and reflective learning logs are good examples of this. It is likely that discipline specific characteristics will emerge here: social presence and reflective practice are often highly valued in teacher development programmes, for example.

Method 2

Both researchers felt that this model produced a deeper focus on activity relating to learning. It did, however, undervalue the discursive aspects of the conversation and social postings were ignored by this approach. In places, the application of this method was difficult to align with the intended learning outcomes – especially where a participant simply expressed agreement with a point of view. In this sense such 'hidden' learning is very difficult to measure.

Again, some overlap of categories became apparent. The separate categories of comprehension and application were neatly spanned by participants' postings that described prior experience (comprehension) in the context of a real world example (application). Should postings such as these be counted in both categories, with the complexities of double counting and the effect on numerical analyses? Should the 'higher level' category take precedence in the analysis?

Method 3

This approach produced the widest variation and least agreement in researcher analysis, apparently due to variation in interpretation of the learning outcomes. One researcher discounted all content that could not be related very closely to the learning outcomes, the other took a broader view and included related content. Both agreed that there were important aspects of discussions that were not explicitly identified using this approach. In particular and in a similar vein to Method 2, this approach does not capture the value of discussion unless it is explicitly cognitive in nature. Moreover, this approach appears excessively narrowly defined in that it excludes any educational benefits of unintended learning outcomes.

Towards a composite method

It is clear from the discussion that all three methods proved to have a number of limitations. Key aspects revolved around the lack of mutually exclusive categories and the degree of focus on the cognitive content. In this respect, Method 2 has particular strengths in terms of the analysis of cognitive content, but uses a larger number of categories which leads to problems with classification. Whilst some interesting findings about student learning were made possible by this study, the problems encountered in implementing each method (in particular the low inter-rater reliability in some parts) give cause for concern. We are therefore proposing to pilot a slightly different method, utilising those aspects of the approaches tested which were most successful.

Although not subjected to statistical analysis, the reliability of the first method appears to be satisfactory. However, the large amount of material categorised as 'cognitive presence' provides little insight into the level and depth of student engagement. The second method (using Bloom's taxonomy) provided more interesting data on student learning but suffered from more substantial reliability problems and discounted all elements of 'social presence'. A composite method, involving selecting the key parts of Bloom's taxonomy and combining them with a measure of social presence, might therefore enable a more meaningful analysis. To this effect we have adopted the revised Taxonomy proposed by Anderson and

Krathwohl (2001) and divided the six levels into two broader subgroups: that of lower level (spanning 'Remembering', 'Understanding' and 'Applying') and that of higher level (spanning 'Analysing', 'Evaluating', and 'Creating').

The categories we propose are therefore:

1. Lower level engagement (Prior learning and experience: Remembering, Understanding, Applying)
2. Higher level engagement (Making inferences and developing new knowledge: Analysing, Evaluating, Creating)
3. Social presence (Indirectly supporting the learning experience)
4. Tutor facilitation
5. Off task discussion

This combined approach will, we hope, provide more detailed measures of cognitive engagement than a collective 'cognitive presence' category, without attempting to resolve this category in overly fine detail. Where postings include an element of social presence and cognitive engagement, we propose to classify the posting as lower or higher level engagement, reserving aspects of postings that are purely social for the category of 'social presence'. Whilst the focus will be on the activities of the learners, tutor postings also need to be quantified (and excluded from total word counts) in order to avoid skewing the results. We also believe it important to quantify the amount of off-task discussion to provide a more accurate measure of 'signal to noise' ratio in this respect.

Concluding remarks

Wider use of these kinds of methodologies provides the potential to enhance e-learning research and evaluation. Benefits include readily available data and the possibility of making comparisons between different types of OAD both within the same course (as in this study) but also between different courses. It would be interesting to note, for example, whether the low level of use of prior knowledge and reading citations were specific to this course, or are a more general feature of online discussions due to their informal nature. In addition, the relatively high level of real-world examples noted in this study may be due to the nature of this course, a practically-orientated and vocational programme of study. Alternatively, this may be indicative of the kind of discourse encouraged by online discussion. These issues are worthy of further study.

References

Anderson, L. W., Krathwohl, D. R. & Bloom, B. S. (2001). *Taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives.* London, Longman.

Bloom, B. S. (1956) *Taxonomy of educational objectives: The classification of educational goals*. *Handbook 1, cognitive domain*. New York: Longmans, Green.

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

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *Internet and Higher Education*, 11(2), 1-14.

Gilbert, P. K., & Dabbagh, N. (2005). How to structure online discussions for meaningful discourse: A case study. *British Journal of Educational Technology*, 36(1), 5-18.

Murphy, E. (2004) Recognising and promoting collaboration in an online asynchronous discussion. *British Journal of Educational Technology*, 35(4), 421-431.

Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks*, 7(1), 68-88.

Volet, S., & Wosnitza, M. (2004). Social affordances and students' engagement in cross-national online learning: An exploratory study. *Journal of Research in International Education* 3, 5-29.

Author contact details

Deborah Cotton, Educational Development and Learning Technologies, University of Plymouth, Plymouth PL4 8AA, UK. Email: debby.cotton@plymouth.ac.uk.
WWW: http://www.plymouth.ac.uk/staff/dcotton.

Jon D. Yorke, Educational Development and Learning Technologies, University of Plymouth, Plymouth PL4 8AA, UK. Email: jon.yorke@plymouth.ac.uk. WWW: http://www.plymouth.ac.uk/staff/jdyorke.