Exploring the quality and effectiveness of online, focused peer discussions using the MOOCchat tool

Carl A Reidsema, Lydia Kavanagh, Emmi Ollila, Stephanie Otte and Julie E McCredden
Faculty of Engineering Architecture and Information Technology, University of Queensland
Corresponding Author Email: j.mccredden@uq.edu.au

CONTEXT

Over the past 3 years, students in the first year common course: Engineering Modelling and Problem Solving, have engaged in collaborative discussions of engineering materials concepts through the use of an online tool called "MOOCchat". The MOOCchat tool is an online variation of the typical *peer instruction* protocol (Mazur, 1997) whereby students are given a multiple choice question to vote on, then are asked to discuss their answer with their peers, and then are given a revote. MOOCchats are used to facilitate online collaborative discussions that can reinforce each student's individual learning as they work through the weekly online modules.

PURPOSE

The aim was to investigate whether the use of the MOOCchat tool would facilitate student understanding of core weekly concepts. To substantiate this investigation, analysis of online discussion data and the development of a scale that could be used to classify chats in terms of the quality of student interactions were required.

APPROACH

The data from two years was used to investigate the effects of MOOCchats on short term learning by counting the numbers of students that shifted to from incorrect to correct answers as a result of the chat. The effects on long term learning were then evaluated by looking at the flow on effects of the MOOCchat shifts on exam marks. To create a scale for classifying chats, a two-step process was used. First, content analysis was performed on a subset of the 2014 chats so as to identify a set of meaningful categories that best described the data. These categories were then transformed into two simpler 'Quality of Chat' scales (Individual Interaction and Group Depth of Collaboration) that were given to the 2015 students to rate the quality of their interactions.

RESULTS

MOOCchat experiences helped about one quarter of the students to shift from incorrect answers prechat to correct postchat. Groups who had at least one member who was correct prechat showed greater gains in the short term. In the long term, students who ended up with more correct answers postchat fared slightly better on the subsequent mid-term exam. This was most true when the concepts were easy or medium, while harder concepts showed fewer learning gains as a result of the chats. The content analysis resulted in a Depth of Collaboration scale with 3 levels: shallow, one-way (just telling or agreeing), and integrative interaction. The derived Quality of Chat scales showed that students rated themselves better than they rated their groups, and also showed benefits for groups whose collaborations were focused on understanding. However, these benefits were reduced when concepts were harder.

CONCLUSIONS

MOOCchat discussions are able to develop student understanding in an online blended learning course if the concept questions and the task goals are appropriately designed.

KEYWORDS

Online discussion, online chats, collaborative learning, critical thinking, MOOCchat.

Introduction

A very large flipped classroom course, ENGG1200, has been running in first year engineering at the University of Queensland for 5 years (Reidsema, Kavanagh, Smith, & McCredden, in press). This course uses online modules for learning of the materials engineering concepts, and face-to face time for project work. Over time, investigations of student reflections have revealed ongoing issues around learning of online content. Students struggle with some of the materials concepts which are difficult and can result in misconceptions (Krause, Decker, and Griffin, 2003). This lack of comprehension affects students' ability to integrate the theory into their subsequent applied project work (see Reidsema, Kavanagh, and McCredden, in current proceedings).

In face-to-face settings, the method of *peer instruction* has been used to promote students' understanding of content (Mazur, 1997). This method uses small group discussions during class time. A concept to be considered is given in the form of a multiple choice question which students first answer individually (often using a clicker response system), and then they are asked to convince their neighbour of their answer. Then, after a few minutes of discussion, students are asked to answer the same question again. Peer instruction has shown consistent improvements in the number of correct responses after the discussion, and in correct answers to concept questions given at the beginning and the end of the course. This method has thus been shown to be an effective method for helping students to learn new concepts and to correct their misconceptions in physics and science (Crouch and Mazur, 2001; Smith et al., 2009).

Online learning environments have allowed researchers to investigate peer interactions more closely, due to the ability to record online text-based discussions, which can be used for later analysis. Using a social and cognitive constructivist framework, Garrison, Anderson, and Archer (2001) have developed a model of online discussions, by describing three interacting components, which have been adopted in this study for exploration of the MOOCchat discussions, as follows. 1. There is a *social presence* component based on purposeful connection i.e. students use the online interaction for achieving a common purpose (Garrison, Anderson, and Archer, 2010). 2. There is a *cognitive presence* component which is used to promote *critical thinking*, comprising four stages: a triggering event (sense of puzzlement), exploration (brainstorming), integration (attempts at solution construction) and resolution (critical evaluation and commitment). 3. There is a *teacher presence* component which is created by facilitation of the discussion by the teacher or by the teacher's a priori design of the online discussion task (which subsequently affects how students interact).

There have been several attempts to develop methods that can facilitate student thinking and problem solving during online interactions (for a review of the recent literature, see He, 2013). One such method is the use of online, small group, focused discussions using an online tool called MOOCchat (Coetzee, Lim, Fox, Hartmann, and Hearst, 2015).

MOOCchat was originally conceived of by Berkeley's MOOCLab Research Team to investigate the benefits of using Computer Mediated Communication to address the low completion rates of MOOC courses (Coetzee et al., 2015). MOOCchat involves online interactions within small groups via text, in real time. The MOOCchat process is a 3 step process (Figure 1), whereby students are asked to (i) individually consider a multiple choice concept question (MCQ1), select an answer and create a written justification, then (ii) show their answers and justifications to one another and discuss together to try to find the right answer as a group, and then (iii) each student to revote on what they now consider to be the correct answer (to MCQ1), and to give their final written justification. The MOOCchat program records the pre- and postchat responses, as well the contents of each chat session, allowing for in-depth investigations into the quality of online small group discussions around concept questions.

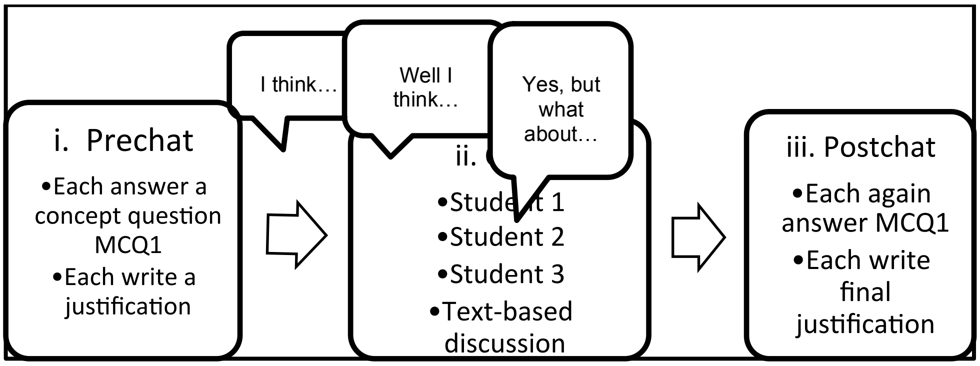


Figure 1: The MOOCchat Process

Coetzee et al. (2015) used data analytics to investigate chats from crowdsourcing worker data, regarding a critical thinking task. The results showed that most group chats (82%) included some form of 'substantive discussion;' which comprised responses to others views, including debates and attempts to persuade. However, no relationship was found between the number of substantive statements and the correctness of answers. Overall, Coetzee et al's results suggest that both the social (purposeful) and cognitive (critical thinking) aspects of online discussions are facilitated by MOOCchat. These findings hold promise for using MOOCchat in a flipped classroom setting so as to as to enhance students' understanding of concepts delivered via online modules. However, to assess the potential learning gains, the benefits of MOOCchat to understanding need to be investigated further.

While Coetzee et al reported high percentages of substantive discussion, this is inconsistent with the bulk of the literature, which has revealed low levels of critical thinking during online interactions. Most of the research has revealed that undergraduates use online discussions more for serial monologues than for interactive debate (Pawan, Paulus, Yalcin, and Chang, 2003). Chats made by science and engineering students during video streaming focused on technical and social-emotional issues, with fewer discussions focused on course content (He, 2013). Furthermore, studies have repeatedly shown that students tend to stay at the exploratory phase rather than moving to the integration and resolutions phases of critical thinking (Garrison et al., 2010; Gunawardena, Lowe, and Anderson, 1997; K. A. Meyer, 2004; Pena-Shaff, Martin, and Gay, 2001). Even when focused on problem solving for project work, students still rarely engage in challenging each other's ideas (Curtis and Lawson, 2001).

The ways in which online group discussions can be enacted so as to elicit deep collaboration and critical thinking is still relatively undeveloped (Garrison et al., 2010). However, short term focused online discussions such as MOOCchat provide new territory for utilising online discussions so as to help facilitate both engagement and understanding (Coetzee et al., 2015). For example, MOOCchat could be used for group discussions of engineering materials concept questions. From the results of (Coetzee et al., 2015) we would expect that the MOOCchat discussions would promote students evaluation of each other's ideas, and that these interactions would help our students to better understand the core course concepts.

A joint collaboration between Berkeley and The University of Queensland began in 2014. The MOOCchat tool was re-purposed to contribute to the learning of concepts inside the weekly online modules, for the first 6 weeks of the ENGG1200 course. The MOOCchat tool has been developed by the eLIPSE centre at the University of Queensland, which has the goal of developing online tools that both improve student learning experiences and at the same time, collect student usage data. Thus the MOOCchat tool recorded all aspects of students' MOOCchat interactions. This data has allowed us to pursue investigations in order to answer the questions below:

1. **Were there any discernible learning gains**? Were the chats able to improve understanding in the short-term (postchat) and the long term (exam marks)?
2. **Were the interactions substantial with regards to the quality of the interaction**? What dimensions of the conversations are important for evaluating the quality of the chats? If these dimensions are used to create scales for evaluation, can these scales then be used by students to self-evaluate the quality of their own interactions?
3. **Are deeper collaborations more beneficial to learning?** Do students and groups who rate more highly on the quality of chat scales have better postchat outcomes?

Investigation 1: Learning gains

Method

MOOCchat conversations were run and recorded for the first 6 weeks of the semester. Each week, two large MOOCchat sessions were available on two different evenings of the week. In each session, the MOOCchat program randomly allocated about three hundred students to triads where they were anonymous to each other, using the identities Student1, Student2, and Student 3. Each triad then followed MOOCchat steps (i) to (iii) described above. The answers that students gave before and after the chat, as well as the contents of each chat were recorded by the MOOCchat program. This allowed for exploration of how each student's ideas changed as a result of their group interactions. For the following analysis, data from weeks 1 to 4 were used, as these weeks were relevant to the subsequent mid-semester exam.

Results

The data from the 2015 cohort were **first** inspected to see how all students fared across weeks 1 to 4. The majority of students gave correct answers postchat for two or more of the four chat sessions (i.e., mean postchat responses correct=2.5, Standard deviation=1.2, N=1,046).

Short term effects

In the short term, MOOCchat group interactions have the potential to change students thinking as a result of the chat; either to help students to learn new facts or to overcome misconceptions related to the specific concept question used for the chat. To investigate this possibility, we coded the effects of each students experience in the chat room into one of four possibilities, according to whether they had the prechat questions correct or incorrect, and then whether they had the answer correct or incorrect postchat, as shown in the first two columns of Table 1.

Table 1: Percent changes in students' answers pre-to postchat

Code	Pre- to Postchat Change in Responses	MOOCchat Session (Week and Session) (N)								
		W1S1 (349)	W1S2 (415)	W2S1 (430)	W2S2 (279)	W3S1 (429)	W3S2 (214)	W4S1 (433)	W4S2 (246)	Av. (%)
1	1: I→I	3.4	27.5	25.3	56.6	5.1	41.1	12.7	30.1	20.7
2	2: C→I	0.9	5.5	2.1	8.6	2.1	6.5	1.8	7.7	3.8
3	3: I→C	12.9	25.8	28.1	30.1	28.7	25.7	30.5	33.7	26.5

4	4: C→C	82.8	41.2	44.4	4.7	64.1	26.6	55	28.5	48.9

The students who received code 3 are of the most interest, as they are the ones who changed their minds in the chat room as a result of the discussion with their peers, and thus moved from incorrect prechat to correct postchat. The students who received code 1 are those who were initially incorrect but still incorrect after the chat, who did not seem to be helped by the chat. Students who received code 4 began and ended correctly; however they may still have benefited by having their understanding consolidated during the chat.

The change codes were used to investigate the before to after transformations that occurred in each session from week 1, session 1 (W1S1), to week 4, session 2 (W4S2), as shown in Table 1. The pattern of results was similar for all sessions, except for W2S2. Closer inspection of this question revealed that that the wording of the question was too hard for students to understand, thus W2S2 was omitted from further analysis. The average percentages across the rest of the sessions for codes 1 to 4 are shown in the final column in Table 1. These averages show that overall, almost half of the students knew the answer prechat and kept the correct answer postchat (code 4), about a quarter started with an incorrect answer but then changed their mind to give a correct answer postchat (code 3), about one fifth started incorrect and stayed incorrect (code 1), and a very small percent moved from a correct to an incorrect answer as a result of the chat (code 2). This pattern of results shows that the chats were effective in helping significant numbers of students to shift to the correct answer.

Number of group members who know the answer

It has been shown that having one student or more in the group who knows the answer can help the others to improve their answers (Coetzee et al., 2015). Using the 2015 cohort, we investigated the week 2 session 1 chats (N= 330) to see whether this occurred for our cohort. The results (Figure 2) showed that irrespective of how many people in the chat group gave the correct answer beforehand (including none) there was an improvement in the number of students who were correct postchat. However, having at least one person who knew the answer beforehand resulted in greater increases in the number correct postchat.

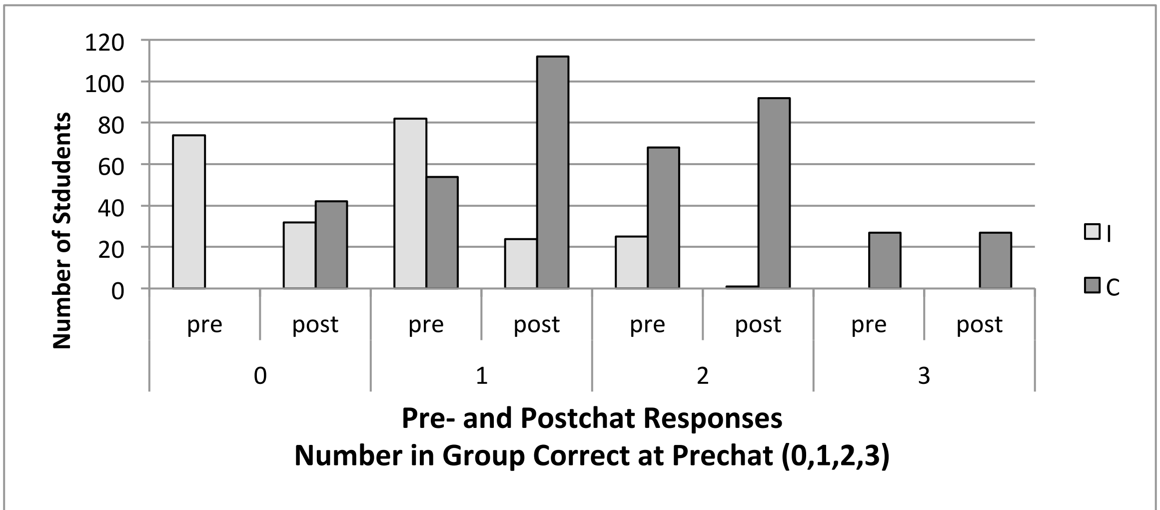


Figure 2: The number of W2S1 students who were incorrect (I) or correct (C) at prechat and at postchat according to how many in their group had the answer correct at prechat.

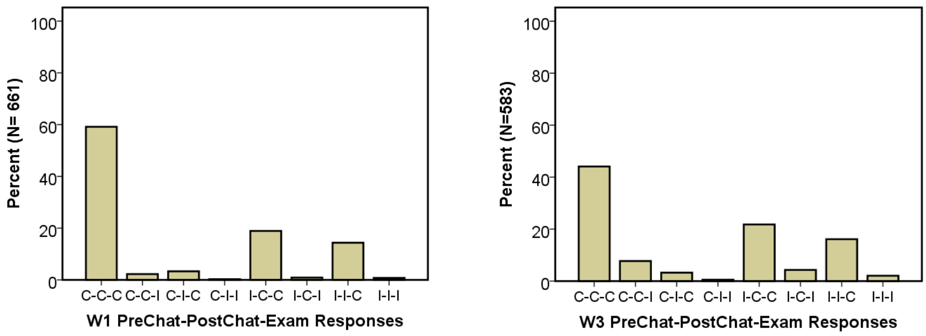
While the discussions inside the MOOCChats helped student to achieve more correct answers postchat, we cannot assume that this means that the students who changed their answers improved their understanding of the concepts. For example, students may have just agreed with the person/s in their chat who knew the answer, rather than learning about the concept. This possibility will be explored in the sections below.

Long Term effects of MOOCchat

If it were the case that the effects of chats was to consolidate or add to correct thinking, or to change incorrect thinking, then it would be expected that those who had more correct outcomes after the chats would show improved understanding of the MOOCchat concepts in a later exam. To investigate this possibility, we looked at the pattern of changes in answers for each chat session compared to the exam marks for the related concept question. The data for the 2015 cohort were used for this investigation (with W2S2 data omitted as before).

Using the number of students who had the prechat questions correct versus incorrect, the difficulty of the concept questions can be ranked as follows: week 1 easy, week 3 medium, week 4 hard and week 2 very hard. This understanding that the patterns of results were different for each week informs the following analysis, where the weekly results are kept separate rather than being collapsed across the weeks.

Each MOOCchat concept question was related to one or more of the exam questions. The pre- to postchat change codes were used to investigate the relationship between the effects of the chat (pre to post) and the mid-semester exam question to which the concept related. Thus, each student's answers to the prechat, postchat and exam questions were classified into one of eight possible sequences that were labeled according to whether answers were correct or incorrect at each stage; e.g., I-C-C was the sequence label for students who answered incorrectly prechat, correctly postchat and then correctly on the related exam question. For each week, the percentages of students in each of the eight prechat-postchat-exam sequences were calculated (Figure3). For example, for those students who completed the week 1 MOOCchat, about 60% were correct prechat, correct postchat, and correct on the later related exam question (sequence C-C-C in Figure 3, W1).



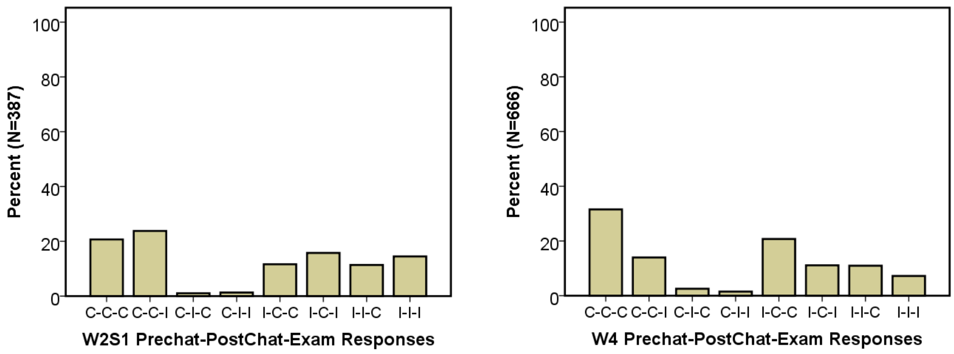


Figure 3: For each MOOCchat session, the percentage of students in each of the prechat-postchat-exam response sequences (C=correct, I = Incorrect).

Weeks 1 and 3 had similar patterns, in that most students were grouped into the sequences C-C-C, I-C-C or I-I-C, which shared the common characteristic of giving correct responses to the exam questions, irrespective of how students fared pre- and postchat. However Weeks 2 and 4 showed different patterns, suggesting that these concepts were harder.

In a study of the benefits of peer instruction face to face, Smith et al. (2009) used peer discussion during class time and then investigated transfer to a similar problem asked later in the class. These authors concluded that that peer learning helped create better understanding, resulting in improved understanding over time. These conclusions were based on an increase in the percentages of correct responses to the later transfer problems over and above the percentage of correct responses before and after the peer discussions. This pattern was consistent for easy, medium and hard questions. However, the above results do not completely mirror the (Smith et al., 2009) findings., because the difficulty of the concept had a mediating effect on the effectiveness of MOOCchat discussions to bring about improved understanding. For example, the week 2 concept was the course threshold concept (J. H. F. Meyer and Land, 2003) *Young's Modulus* (elastic modulus), which is a complex concept relating two component concepts to one another; i.e., *stress* (Force / cross sectional area) and *strain* (change in length / initial length), both of which are also new concepts for the students. While some students benefited from the chats by increasing their likelihood of a correct response postchat (Figure2), students' long term understanding of this concept did not benefit markedly from peer discussion. A greater proportion of the students answered the exam question incorrectly, including those who had the MOOCchat question correct at postchat (Figure 3). This finding has very real implications for teaching, suggesting that peer instruction alone may help, but will not suffice, for ensuring students' understanding of difficult concepts.

There are several possible reasons for why discussing concept questions with peers had benefits for some students but not for others. The students who did not benefit could have just been less able to grasp the concepts, especially if they were difficult, or it could be that their chat experience was less than helpful in promoting understanding. To understand when and how the interactions in the chat room were able to help to improve understanding, we needed to explore the chat sessions in more depth, as described below.

Investigation 2: Quality of Collaboration Scales

MOOCchat outcomes can be based on students' desires to get the answer correct rather than to understand. For example, consider the two different chats below. In the first example, the group is aiming for correct answers, whereas in the second, there is an honest attempt to understand, which results in a new, shared understanding for the whole group.

Example 1: Shallow interaction. The justifications of each student before and after the chat are correct, but unchanged, and do not explore the component concepts of stress and strain.

Written Justifications (Prechat and Postchat)

(Student 1): (E), Justification: Young's modulus is constant for a material and by changing the length or cross sectional area this will not impact on the modulus
(Student 2): (E), Justification: Because the elastic modulus is a property of the material.
(Student 3): (E), Justification: Elastic modulus in based on the type of material not the size or shape.

Chatroom

(Student 3): cool, we all agree
(Student 2): yeah, seems like it, anyone not know this?
(Student 1): it was a pretty easy question

Example 2: Thoughtful consideration of each other's ideas. Integrative interaction.

(Student 2): if the question is asking for the increase in elastic modulus, that is definitely the ratio of stress to strain
(Student 2): so thats why I thought it was c
(Student 3): Well, materials always have the same E value, regardless of the stress/strain conditions. If you increase the cross sectional area, you decrease the stress, but the strain will increase proportionately
(Student 1): ^^ beat me to it
(Student 2): oh ok
(Student 1): I'm leaning towards E now, the Modulus of Elasticity is specific to the material in question. So if the question was what will change the modulus then E is the only option as the others don't change the material properties
(Student 2): so its definitely constant
…

Observations of interactions in the chat rooms such as these could be used to discover the important *quality of collaboration* dimensions, which can most effectively describe what Stahl (2009) has termed 'collaborative moments' in group discussions. It was expected that insight into these dimensions would help to qualify the types of MOOCchat interactions able to promote improvements in students' understanding.

Method

One way of establishing a set of dimensions that can be used to describe the quality of interactions is to employ *content analysis*. This method uses qualitative analysis to explore transcripts so as to find examples of pre-existing theoretical categories, or to find themes which emerge naturally from the data, or both (Garrison and Cleveland-Innes, 2005; Gunawardena et al., 1997; Henri, 1992; K. A. Meyer, 2004; Newman, Johnson, Webb, and Cochrane, 1997; Pena-Shaff et al., 2001; Rourke, Anderson, Garrison, and Archer, 2001; Zhu, 1996). For example, Pena-Shaff et al. (2001) discovered emergent categories such as: *question, reply, support, consensus building (conflict, building arguments together, generating shared conclusions), clarification / elaboration, and social interaction*.

Given the newness of the MOOCchat protocol, we chose to use a mixed (thematic plus emergent) approach to content analysis. The week 2 chats were used for an in-depth investigation of the quality of the chats, because week 2 discussions were more likely to show deeper interactions as a result of students struggling to relate the Young's Modulus component concepts of *stress* and *strain* to one another (as described above). Within the week 2 data, 28 chats from 84 students (16% of the session 1 students) were investigated intensively, using the 2014 cohort (as this data was available when the investigation began). These chats covered a wide variety of interactions of varying length.

The content analysis began with a large set of predefined possible themes, each with their own levels, such as how much students tried to listen to and consider others ideas, how much students tried to explain their own ideas, and how much students struggled with ideas. As the analysis proceeded, it became clear that most of the themes were encompassed within one overarching theme of *Depth of Collaboration* that had three main groupings:

1. **Shallow**: No relevant discussion / No comments about the matter/ No real discussion needed according to students' view (see Example 1 above).
2. **One-Way / Medium**: Telling / Just agreeing / Asking: Some interaction: Door opener / ask others for their ideas (what do you think) / ask others to justify their ideas (why) / agreeing with others; / State own ideas / Express opinion / Disagree with peer without justifying / Justify own answer/ Affect (convince a peer without listening to them).
3. **Integrative**: Co-operative chat / Debate answer/ Disagree with a peer and explain why peer is wrong / Reason with a peer / Clarify something self or peer has said /Ask for justification or clarification explaining why this is needed / Ask a deep question from peer / Explain to a peer on a deep level / Express a deep point that is relevant to discussion / Be active in chat and strongly aim to learn from peer (see Example 2 above).

The long term aim for this project is to create a scale that can be used for assessment purposes, as well as for further analysis; e.g., to be able to relate the depth of collaboration to assessment outcomes (Arbaugh et al., 2008). Thus the *Depth of Collaboration* scale will be developed further in future studies to allow for more rigorous analysis of student interactions.

Student-friendly scales: In the short-term, the aim was to create a student-friendly scale informed by the Depth of Collaboration scale that could be used by students to rate themselves on how well they had been collaborating in the chat room. Pawan et al. (2003) have recommended asking students to self-code their contributions to discussions, to help them to develop metacognitions and thus deeper collaborative skills (Duffy, Dueber, and Hawley, 1998). Given the findings described above that undergraduate students are not good at critical thinking during online interactions, and given that our students were novice first years, any scale would need to be simple to use yet able to provoke students' thinking about their group interactions. Thus the Depth of Collaboration scales were adjusted slightly in order to derive two student-friendly scales that could help to highlight for the students the importance of their individual interactions and their group's interaction, and of how the former contributes to the latter. These two scales and their relationship to the original Depth of Collaboration scale are shown in Table 2.

Table 2: The two Quality of Chat scales derived from the Depth of Collaboration scale

<table>
<tr><td>Depth of Collaboration</td><td colspan="2">Quality of Individual Interaction Scale</td><td colspan="2">Group Depth of Collaboration Scale</td></tr>
<tr><td>Shallow</td><td>1</td><td>I gave my own ideas but did not really consider others ideas</td><td>1</td><td>All stated their own ideas but with no real interaction</td></tr>
<tr><td rowspan="2">One-way/Medium</td><td rowspan="2">2</td><td rowspan="2">I considered others ideas but did not really state my own</td><td>2</td><td>Mostly just agreed with the person/s who seem to know the answer</td></tr>
<tr><td>3</td><td>Discussed ideas to get the correct answer but not to fully understand</td></tr>
<tr><td>Integrative</td><td>3</td><td>I gave my own ideas and I considered others ideas</td><td>4</td><td>Discussed ideas to get the correct answer and to fully understand</td></tr>
</table>

The Quality of Chat scales were then used with the student cohort in the following year (2015). After each MOOCchat session, students were given a postchat survey which included the two Quality of Chat scales. The ratings that students gave themselves and their group on each scale for each week were summed across the 4 weeks (Figure 4).

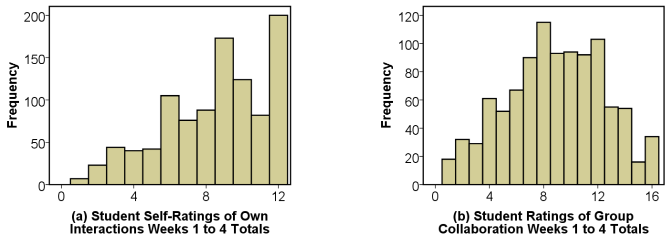
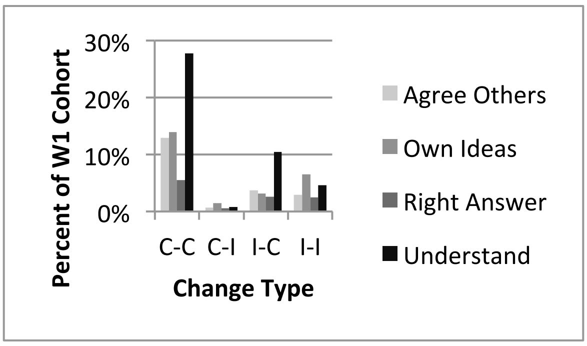


Figure 4: Students' ratings for (a) Quality of their individual interactions and (b) Depth of their group's collaboration totalled across weeks 1 to 4.

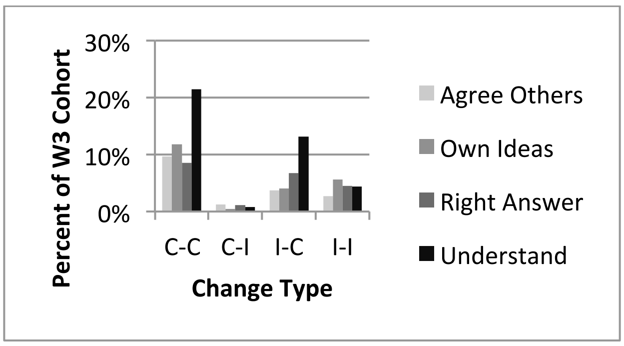
Comparison of the frequency distributions (Figure 4 (a) vs 4(b)) shows that approximately 200 students rated themselves highly but rated their group as medium interactors. This discrepancy could have been overrating of self, due to students either not being able to be objective or having little understanding of what constitutes 'consideration of others ideas'. It could also have been that the group dynamics were such that some students wanted to discuss while others did not. The Group Depth of Collaboration scale was thus chosen for further analysis due to the likelihood of those ratings being more objective and valid.

Quality of Collaboration and Learning

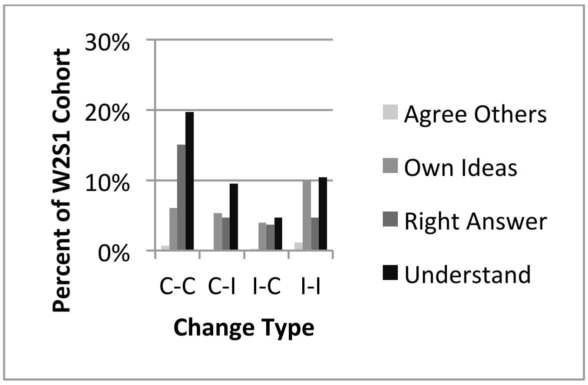
The Group Depth of Collaboration scale was used to explore the relationship between the quality of each group's chat and the changes in responses that members made as a result of the chat. The results for weeks 1 to 4 (Figure 5) show that the students who were correct pre-chat and who ended correctly postchat (C-C) were in groups that mostly aimed for *understanding*. This is true for all weeks. Similarly, for weeks 1 and 3, the students who were incorrect prechat but who ended correctly postchat (I-C) were more likely to be in groups focused on understanding. Thus, for easy and medium questions, group discussions that focused on understanding helped students to shift to the right answer. However, Figure 5 also shows that when the concepts were hard (week 4), the students who were correct postchat (C-C or I-C) were likely to be in a group that just agreed with person who seemed to know the answer. Furthermore, when the question was very hard (week 2), students who ended up correct (C-C or I-C) were in groups that often just focused on getting the right answer rather than understanding.

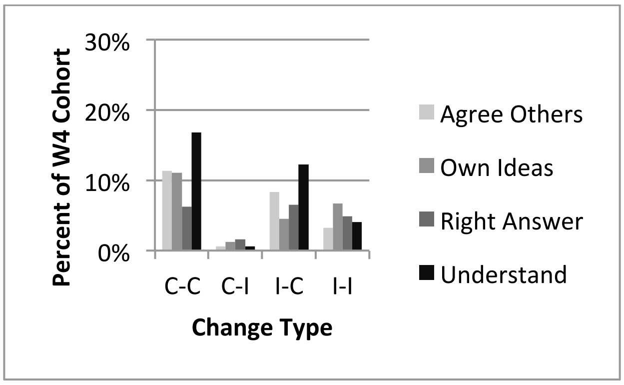


W1 Easy concept questions



W3 Medium concept questions





W2 Very hard concept question **W4 Hard concept questions**

Figure 5: The different levels of group collaboration related to members changes in response pre- to postchat for week 1 to 4 MOOCchats.

The week which had the hardest question (week 2) had the highest percentage (10%) of groups where members were incorrect prechat and remained incorrect postchat. Some of these students aimed for understanding, while some just stated their own ideas and did not try to interact. Surprisingly, quite a few (10%) of the week 2 students moved from correct prechat to incorrect postchat, even though they were in groups that aimed for understanding. Overall, it appears that groups that had understanding as their goal helped their members to end up with correct responses, except for when the concepts being discussed were very hard. In this case, aiming for shared understanding caused many of the students to end up incorrect at postchat. Thus online peer discussion does not guarantee improved understanding when concepts become more difficult.

Summary

The above investigations have shown that online discussions focused on problem solving such as those facilitated by the MOOCchat tool can help students to shift from incorrect to correct responses. The learning gains are greater when the concepts being probed are of an easy or medium level of difficulty, or if other group members know the answer at prechat. Such improvement in answers due to the MOOCchat discussions seems to indicate a shift in understanding that endures it the long term (i.e. when tested in exams). However, when the concepts being probed are hard, the gains in students' understanding due to the chat are lower and less enduring. The quality of the group discussions also has a mediating effect on learning outcomes. The MOOCchat discussions provided many examples of integrative collaboration (see Example 2 above), which reveals the tools potential in facilitating critical thinking over and above other tools described in the literature. Groups that aimed for deeper understanding helped their members to move from incorrect to correct responses. However, when concepts become more difficult and if the students are struggling with the concepts, aiming for understanding is less likely to have positive outcomes from the MOOCchat sessions.

These above results and insights have implications for teachers using blended learning tools. Online peer discussions can help to supplement student's study of online content, so as to improve understanding. However the concept questions need to be designed to sit within the student's 'zone of proximal development' (Vygotsky, 1978) so that peers are able help each other to construct shared understanding. If the concepts are hard, a greater teacher presence is likely to be needed. The method of introducing a mini-lesson in to the MOOCchat (Coetzee et al., 2015) or use of online help tools such as Caper Q and A (Herbert, Smith, Reidsema, and Kavanagh, 2013) are ways in which this may possibly be achieved in a flipped classroom setting. Furthermore, teachers need to encourage students to focus their goals for online discussions towards understanding rather than correctness. Reflection on self-contributions may be promoted by the use of postchat ratings such as the above Depth of Group Collaboration scale. Allocating marks for students' quality of interactions may also motivate students to engage in deeper interactions. However they may still need help to overcome blockages to critical discourse such as interpreting criticisms as personal attacks (Rourke and Kanuka, 2007).

Several directions for future research are suggested from the above results and analysis. The mediating effects of concept difficulty as well as the different goals of group members confirms that all three components of critical thinking: social, cognitive and teacher presence are present and interacting within the online MOOCchat environment, as suggested by (Garrison et al., 2010). Furthermore, using correctness of answers as a measure of learning gains does not give full insight into the effectiveness of peer discussions. Developing scales for rating the quality of chats that can be used by students as well as researchers is one way of achieving greater insights. The MOOCchat tool provides a user-friendly environment for developing such scales.

Conclusions

The MOOCchat environment promotes online peer interactions that can help students to gain understanding. However, the difficulty of concepts is a mediating factor that needs to be taken into consideration when designing concept questions. A student-friendly scale for self-rating of the quality of their MOOCchat interactions was developed. Use of this scale revealed that the quality of collaboration during the chats was also a mediating factor in learning gains. Future MOOCchat tasks need to be designed into the course so as to focus student interactions towards understanding and integrative collaboration. Further studies will develop the concept questions and the task environment so as to achieve these aims.

References

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework using a multi-institutional sample. *The internet and higher education, 11*(3), 133-136.

Coetzee, D., Lim, S., Fox, A., Hartmann, B., & Hearst, M. A. (2015). *Structuring interactions for large-scale synchronous peer learning.* Paper presented at the Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American journal of Physics, 69*(9), 970-977.

Curtis, D. D., & Lawson, M. J. (2001). Exploring collaborative online learning. *Journal of Asynchronous Learning Networks, 5*(1), 21-34.

Duffy, T. M., Dueber, B., & Hawley, C. L. (1998). *Critical thinking in a distributed environment: A pedagogical base for the design of conferencing*. Retrieved from

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), 7-23.

Garrison, D. R., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The internet and higher education, 13*(1), 5-9.

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *The American Journal of Distance Education, 19*(3), 133-148.

Gunawardena, C. N., Lowe, C. A., & 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), 397-431.

He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior, 29*(1), 90-102.

Henri, F. (1992). Computer conferencing and content analysis *Collaborative learning through computer conferencing* (pp. 117-136): Springer.

Herbert, J., Smith, E., Reidsema, C. A., & Kavanagh, L. (2013). *Helping Students find ansers: Algorithmic interpretation of student questions.* Paper presented at the The Australasian Association of Engineering Education, Gold Coast, Australia.

Krause, S., Decker, J. C., & Griffin, R. (2003). *Using a materials concept inventory to assess conceptual gain in introductory materials engineering courses.* Paper presented at the Frontiers in Education, 2003. FIE 2003 33rd Annual.
Mazur, E. (1997). *Peer instruction*: Upper Saddle River, NJ: Prentice Hall.
Meyer, J. H. F., & Land, R. (2003). *Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines*: University of Edinburgh Edinburgh.
Meyer, K. A. (2004). Evaluating online discussions: Four different frames of analysis. *Journal of Asynchronous Learning Networks, 8*(2), 101-114.
Newman, D., Johnson, C., Webb, B., & Cochrane, C. (1997). Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information science, 48*(6), 484-495.
Pawan, F., Paulus, T. M., Yalcin, S., & Chang, C.-F. (2003). Online learning: Patterns of engagement and interaction among in-service teachers. *Language Learning & Technology, 7*(3), 119-140.
Pena-Shaff, J., Martin, W., & Gay, G. (2001). An epistemological framework for analyzing student interactions in computer mediated communication environments. *Journal of interactive learning research, 12*(1), 41.
Reidsema, C. A., Kavanagh, L., & McCredden, J. E. (2016). *Project design and scaffolding for realising practitioner learning in a large first year flipped classroom course.* Paper presented at the Annual Conference of the Australasian Association for Engineering Education, Coffs Harbour, Australia.
Reidsema, C. A., Kavanagh, L., Smith, N., & McCredden, J. E. (in press). Designing an active learning environment architecture within a flipped classroom for developing first year student engineers. In C. Reidsema, L. Kavanagh, R. Hadgraft, & N. Smith (Eds.), *Flipping the Classroom – Reimagining Practice and Practices* Springer.
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 (IJAIED), 12*, 8-22.
Rourke, L., & Kanuka, H. (2007). Barriers to online critical discourse. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 105-126.
Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *science, 323*(5910), 122-124.
Stahl, G. (2009). The VMT vision. In G. Stahl (Ed.), *Studying virtual math teams* (pp. 17-29). New York: Springer.
Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. Cambridge: Harvard University Press.
Zhu, E. (1996). Meaning Negotiation, Knowledge Construction, and Mentoring in a Distance Learning Course.