

Being social or social learning?

A sociocultural analysis of the FutureLearn MOOC platform

Philip Tubman
Learning Technology
Lancaster University
Lancaster, UK
p.tubman@lancaster.ac.uk

Dr. Murat Oztok
Dept. of Educational Research
Lancaster University
Lancaster, UK
m.oztok@lancaster.ac.uk

Dr. Phil Benachour
School of Computing and
Communications
Lancaster University
Lancaster, UK
p.benachour@lancaster.ac.uk

Abstract: MOOC environments seem to offer the potential for massive amounts of social learning. The numbers of registered learners can be very impressive, but it is unclear whether they realize this potential. MOOC environments have unique challenges for pedagogy which are not seen in other socio-constructivist learning environments. These are the scale of participation and diversity of participation. Learners have access to a wide range of points of view, but few means to filter this, creating information and collaboration overload issues. This study quantitatively examines interaction data from 10 MOOC instances of 6 distinct courses, hosted on the FutureLearn platform. The interaction data is taken as an indicator for depth of learning or knowledge construction in the sociocultural sense. Although levels of participation are high within these MOOCs, most conversations are seen to have surface level interactions. New platform and pedagogy affordances are suggested that may help deal with the ‘overload’ issues which result in these interaction problems. These affordances aim to give more agency to learners, enabling them to be more active in seeking relevant content and interacting with it.

Keywords: MOOC, Social Learning, Vygotsky, End User Development Environments

I. INTRODUCTION

The sociocultural perspective of learning is grounded in Vygotsky's idea that higher level knowledge is necessarily socially created [1]. That is to say “human thinking is characterised by the dynamic relationship between social interaction and individual cognition” [2]. According to Vygotsky, functions of development will appear twice: firstly in the ‘intermental’ where ideas are negotiated socially with a more experienced other, then in the ‘intramental’, when the ideas are internalised and relationships established in conceptual maps. A key concept is the notion of ‘agency’, which is how we develop a sense of self through participation; in other words, the process of ‘situating’ new knowledge. In this sense, participation changes our understanding of our own self [3] and participation is the process through which individuals learn. Human, cultural or technological factors can act as barriers or enablers for participation, therefore restricting or enabling agency and development. “Today's most significant phenomenon in information and communication technology exemplifying agency is perhaps the social media, or ‘Web 2.0.’” [4].

Lapadat argues that the nature of participating in interactive forums can achieve good learning experiences, because “as writing composition typically demands higher order thinking process, there is great potential for conceptual change” [5]. This can only be achieved when the appropriate expectations for participation are set. Positive outcomes will not automatically occur by including a forum in a course design; the socio-constructivist learning model demands high quality participation from learners.

Barricelli et al. [6] introduce the concepts of ‘information, participation and collaboration overload’ in end user development (EUD) environments which support ‘cultures of participation’, a term used by Fischer [7]. They claim that developments are needed in EUD environments to reduce these overload problems. This is important for the design of MOOC environments, where sociocultural learning is the intended outcome, and end user participation is the means of achieving it.

Despite a growing number of summative attempts to understand peer interaction and user generated content through social network analysis [8], content analysis using structural topic modeling [9], and analysis of discussion topic lifecycles [10], more work needs to be done to understand how the MOOC platforms can relinquish more agency to learners themselves, in the sense that they are able to find the ‘right’ others with whom they can interact, participate and ultimately learn alongside.

The purpose of this paper is to explore the present situation with regards peer to peer interaction in MOOCs, by analyzing the learner data from 10 instances of 6 courses developed on the FutureLearn platform by a University in the North West of England. An instance is a unique course event, repeated often, with little or no modification to published content. This is a quantitative study of the user generated data (user comments), examining the interaction between active participants, which is used as an indicator for sociocultural discourse and as a means to analyze the agency that the platform allows its users to have.

The aim is to follow this up with qualitative content analysis of the data, to explore in more detail what types of interactions occur within the constraints of the platform functionality. This will help understand what future developments are needed in terms of platform design, and the parameters of the resulting trade-offs.

II. PAPER ORGANISATION

The paper is organized into a theoretical framework, which examines previous models and instruments for learning generated content analysis. The model which best demonstrates the interactive nature of the writing in MOOCs is chosen. The methodology describes how the chosen model is implemented by taking the FutureLearn platform affordances into account. In the results section, the whole dataset is shown and then data from a specific MOOC is analyzed in detail which brings out specific areas for discussion. The discussion section expands on these and the conclusion suggests areas of further research which would enhance this data or deliver further proofs.

III. THEORETICAL FRAMEWORK

De Wever et al. [11] detail 15 instruments for content analysis study: ‘social-constructivism’, ‘community of inquiry’ and ‘knowledge construction’ are the most common theoretical frameworks. Community of inquiry [12] divides all content into 3 ‘presences’: social, cognitive or teaching. Knowledge construction [13] divides user generated content into 5 ‘levels’, from sharing experience to synthesizing previously stated points. In addition to these, Lapadat [5] states that positive outcomes can only occur when expectations are set as to the quality and quantity of individual contributions as interactive writing on socio-constructivist learning environments demand high quality participation.

MOOC platforms must meet the dual challenges of low barriers to access meaning different ‘levels’ of content, and also scale. This means there is a lot of diverse opinions, which is a good thing in terms of possibilities for synthesis, but this brings with it esoteric responses, which may not be relevant or useful. Filters and notifications need to be developed in order that users can keep up with and make best use of others’ contributions. Without these platform developments, the sociocultural learning process becomes more about serendipity and luck, more akin to learning in a random crowd of people than a structured learning experience.

Due to these challenges, this paper follows the literature back to Henri’s seminal paper on content analysis [14], which divides all discourse into 5 dimensions which describe the holistic nature of sociocultural learning. These are participative, interactive, social, cognitive and metacognitive. Whilst each of these can further break down into sub-dimensions, e.g. interactive could break down into direct, indirect, commentary or response, this sub-division would require qualitative analysis which is out of the scope of this paper. This paper will focus solely on the interactive dimension as an indicator of sociocultural learning through learners’ interactive participation.

IV. METHODOLOGY

This study analyses 6 distinct MOOC courses across a range of subject areas. Several have been repeated several times totaling 10 instances. 39193 learners have joined these courses, and 18991 can be described as ‘social’, meaning that they have made at least one comment on the platform. The platform presents content, such as video lectures, into

consecutive web pages, called ‘steps’, and there is a space to comment on each ‘step’. Learners can make an unlimited number of comments or replies per step. The courses in the study range from 2 weeks to 8 weeks in length with between 9 and 52 steps per week. Each step represents a space where sociocultural learning could occur.

Learners can choose to ‘like’ comments and there are tabs on the page to filter ‘most liked’ comments and also to the ‘replies’ to conversations you are involved in. A single daily email also notifies you of replies to all your conversations. Users can ‘follow’ each other and filter the comments across all the pages to just those ‘followed’ users. Comments are displayed in reverse chronological order, 100 per page, with the most recent appearing at the top and a link at the bottom of the comment stream to view ‘older’ which is the next 100 comments in reverse chronological order. The features described above represent the feature-set through which learners mediate their participation and interactions.

Interaction is analyzed across all 10 MOOC instances to indicate trends of behavior that are occur within the platform. The ‘interaction index’ in this sense takes the length of the conversations as indicative of the depth of knowledge construction. Further qualitative analysis of the data would reveal a deeper understanding of the sociocultural nature and knowledge construction within conversations.

Interactive writing is a key component for learning in computer mediated communication environments [5]. In terms of interaction on the FutureLearn platform, this relates to the length of a conversation, or the count of replies to a comment. It is assumed that replies are directly related to their parent post, so posts with many replies indicate a continuous conversation, which may mean knowledge construction in the sociocultural sense. Qualitative analysis would be required to verify this, but the scope of this paper is to identify trends. It is fair to assume that comments with no replies at all cannot support claims for providing sociocultural learning experiences as they are not conversations or ‘talk’ in the sociocultural sense. This is not to say that it is not possible to learn from reading these comments, but just that they are not interactive in nature, so cannot follow the sociocultural perspective grounded in Vygotsky. These results cannot be taken on their own as firm evidence for or against the presence of sociocultural learning, but as a whole they give an indication of the learner trends on the platform.

V. RESULTS

The total number of comments which are replies (in table 1) are further broken down into their composite conversations by linking each ‘reply’ with it’s parent ID. This gives the length of the conversation/ thread (fig. 1). In theory this could be of an unlimited length. The maximum conversation included 83 replies and is found in the English Literature MOOC about William Wordsworth, a 18th century romantic poet.

No MOOC contained more than 9 unique instances of conversations that exceed 10 replies in length, making up less than 1% of total conversations. Fig. 1 looks at the trends within conversations. Fig. 2 is conversation length with 1-10 replies which is 99% of the total conversations.

TABLE I.

Comments, replies and conversations across all MOOCs

| MOOC name | Total comments | Original posts | Total replies | % of total comments which are replies | Unique threads |
|----------------------|----------------|----------------|---------------|---------------------------------------|----------------|
| Corpus Linguistics 1 | 20046 | 10041 | 10005 | 49.91 | 4127 |
| Corpus Linguistics 2 | 19556 | 12547 | 7009 | 35.84 | 3590 |
| Corpus Linguistics 3 | 9600 | 6225 | 3375 | 35.16 | 1702 |
| Food Security 1 | 20595 | 14956 | 5639 | 27.38 | 3202 |
| Food Security 2 | 18822 | 12790 | 6032 | 32.05 | 2834 |
| Dyslexia | 44152 | 35638 | 8514 | 19.28 | 4893 |
| Ebola 1 | 4892 | 3933 | 959 | 19.6 | 514 |
| Ebola 2 | 1174 | 980 | 194 | 16.52 | 110 |
| Soils | 14347 | 10237 | 4110 | 28.65 | 1830 |
| William Wordsworth | 26156 | 18162 | 7994 | 30.56 | 3101 |

As stated previously, the sample of MOOCs is distributed across all subject areas, from Environmental Sciences and Biological Sciences to English Literature and Linguistics. It is interesting to see in fig. 1 and fig. 2 that the count of replies for all these MOOCs is heavily weighted towards a single reply and only a nominal number of conversations are greater than 5 replies in length.

Given the spread of interactions is even throughout all the MOOCs, this paper will focus in on the Dyslexia MOOC to examine the results in more detail.

The Dyslexia course was 4 weeks long and started in April 2015. 10295 learners started the course, and 5824 learners are described as ‘social’, which means that they made at least one comment. These 5824 learners made 44152 comments overall, of which 8514 were replies to another comment, approximately 20% of the total. This means that 35638

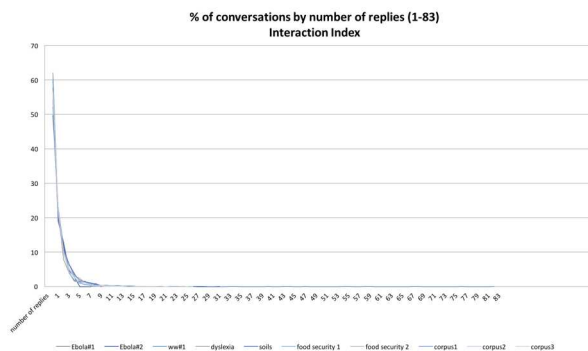


Fig. 1: Percentage of total conversations by number of replies

TABLE II.

Conversation length within Dyslexia MOOC

| No of replies | Count of conversations | % of conversations |
|---------------|------------------------|--------------------|
| 1 | 3032 | 61.96 |
| 2 | 1071 | 21.88 |
| 3 | 383 | 7.82 |
| 4 | 193 | 3.94 |
| 5 | 91 | 1.85 |
| 6 | 46 | 0.94 |
| 7 | 30 | 0.61 |
| 8 | 16 | 0.32 |
| 9 | 9 | 0.18 |
| 10 | 8 | 0.16 |
| 11 | 5 | 0.1 |
| 12 | 2 | 0.04 |
| 13 | 1 | 0.02 |
| 14 | 1 | 0.02 |
| 15 | 1 | 0.02 |
| 16 | 1 | 0.02 |
| 17 | 2 | 0.04 |
| 18 | 1 | 0.02 |

comments (or approx. 80%) were original posts, which may or may not have replies. When the replies are aggregated into conversations, or threads (i.e. replies which share the same parent ID/ original post), there were 4893 unique conversations. It is assumed that the length of these conversations is indicative of the depth of the sociocultural learning occurring and we can already see that the majority of the original posts do not ever receive a reply; in fact, only 4893, or around 15% of original posts do evolve into conversations. Table 2 shows a count and percentage of

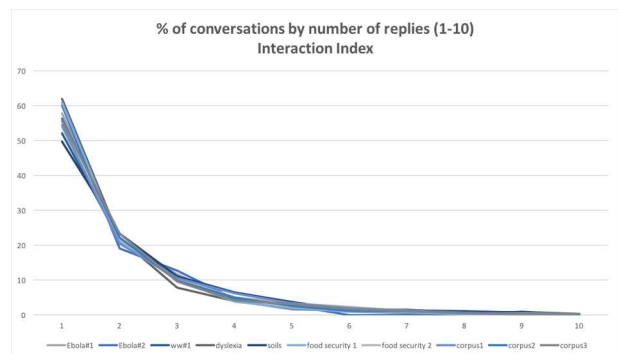


Fig. 2 percentage of total conversations by number of replies from 1-10

conversations for the Dyslexia MOOC in terms of its length. It can be seen that only 2.5% of conversations on the MOOC receive 6 or more replies.

This indicates that sociocultural learning, as defined by high quality interactive writing is likely to be fairly low. That is to say if 61% of conversations consist of a comment and a single reply, it is hard to imagine how they can be synthesizing previously made points of view. It is more likely that these replies express agreement or disagreement with the original post, with no further discussion. Furthermore, it is worth considering the features of the platform when analyzing these results. The platform has controls to filter comments by most liked, replies and people I have followed. You must find and join a conversation to get any notification of updates to that conversation (either by email or by using the platform itself). This may explain why around 84% of comments have fewer than 2 replies, which may be the first reply and the person who made the original post replying again to that. It more difficult for additional learners to join the conversation after a period of time when it is pushed down the comment stream, although if they do join the conversation they will receive a daily notification of updates to it. Further research is needed to establish how many people are involved in the longer conversation threads to establish whether it trends towards being limited to 2 (the original poster and the first replier), or whether more people get involved to share their opinions, and also when they are able to get involved. For example, if more people are involved, are they making their replies at around the same time as the original posts, or are they actively seeking those conversations by manually reading many comments? More qualitative analysis is also required to validate these indicative trends and properly assign the conversation threads to appropriate levels of knowledge construction.

VI. DISCUSSION

Any online course is made up of 3 overlapping factors: the platform features, which determine what users can do, pedagogical features, which is the expectations that the educators set regarding what learners should do, and then the learners themselves, and what they choose to do.

The results of this study indicate that the platform features are the biggest factor for the low level interactions, which would imply a more surface level of learning. Given that 5824 learners chose to 'be social' in the Dyslexia course, and 18991 across the 10 MOOCs, it seems unlikely that they all actively chose to have short conversations. It is more likely that this is to do with either the expectations or the platform features.

As the notification affordance on the platform only informs active participants in a conversation about any replies to it, unless they are also being followed, or it is liked by many others, this may actually restrict the possibilities for conversation depth, and knowledge construction. This is because the only other way of encountering the conversation would be for learners who are on the page at the same time to see the comment near the top of the comment stream or do read all comments on all pages. This is quite limiting and it points towards the need for developing keyword searching for conversations that may have been left some time before and hence is further down the comment stream.

There is a discussion space on each step, and therefore many opportunities for conversation; this may point to the need for a different kind of pedagogical approach when setting expectations for sociocultural learning. Learners clearly enjoy being social on the platform, as the lack of long conversations does not deter learners from leaving comments. However, a single step does not equate to a learning outcome, so learners may not know where it is appropriate to comment and where it is appropriate to discuss. It may be necessary to develop explicit 'discussion steps' at the end of a learning outcome, in which the educators ask the learners to think about the whole learning outcome and engage in meaningful discussion with their peers. This would support the conclusions of Lapadat [5] that expectations for participation need to be clear to create opportunities for written interaction that can support conceptual changes. Further research is needed to establish whether discussion steps made in this way would elicit deeper sociocultural learning, and counter the 'information overload effect' that was identified by Barricelli et al. [6] in all end user development environments.

VII. CONCLUSION

MOOCs are different from other open educational resources in that they are course events. This affords them far more potential as sociocultural learning environments. However, they differ from distance learning programs because of their openness and scale. The platforms are still in their infancy and have yet to fully tackle some of the specific challenges relating to this. They cannot be directly compared to socio-constructivist learning experiences which have come before them, for example in distance and network learning, because they have unique challenges in the scale of participation and managing the diversity of opinion which comes from the courses being open. A balance needs to be found in terms of the levels of notifications for learners; too many notifications and learners will have to deal with participation and collaboration overload issues; too few impairs the interactive potential of the platform.

Currently, learner participation remains unfiltered, and notifications are used conservatively, which is presumably to prevent putting off learners with too much information. This paper shows that this does seem to directly effect the depth of the interactions, and interactive trends suggest only surface level sociocultural learning. Further research is needed to establish the number of people involved in conversations. This would indicate if the combination of the reverse-chronological comment stream and notifications to only active participants in a conversation has and impact on interactivity. Interactivity combined with multiple diverse points of view hold the most exciting potential for massive pedagogy in terms of sociocultural learning; that is the dynamic between social discourse and individual cognition, as described by Mercer [2], but on a massive scale.

There are certainly lots of untapped potential in MOOC learning; in no other space can thousands of interested global learners come together for such an event; to share experiences, construct new meaning and form new social relationships and networks. However, more research is needed to understand how learners can use each other most effectively to fully

maximize this potential. This is both in terms of developing tools for the online learning environment and in terms of the types of pedagogy which is suitable for massive participation.

A combination of platform features and pedagogical modifications is suggested to counter these challenges, the aim of modifications should be to give more agency to learners. MOOCs are informal learning experiences so more work is needed to establish how to support self-directed learning on a massive scale.

REFERENCES

- [1] L. S. Vygotsky, *Mind in society: the development of higher psychological processes*. Cambridge: Harvard University Press, 1978.
- [2] K. Littleton and N. Mercer, *Interthinking: Putting Talk to Work*. Routledge, 2013.
- [3] E. Wenger, *Communities of practice: Learning, meaning, and identity*. Cambridge university press, 1999.
- [4] K. Kumpulainen, L. Krokfors, L. Lipponen, V. Tissari, J. Hilppö, and A. Rajala, 'Learning Bridges – Toward Participatory Learning Environments', Dec. 2009.
- [5] J. C. Lapadat, 'Written Interaction: A Key Component in Online Learning', *J. Comput.-Mediat. Commun.*, vol. 7, no. 4, pp. 0–0, Jun. 2006.
- [6] B. R. Barricelli, G. Fischer, A. Mørch, A. Piccinno, and S. Valtolina, 'Cultures of Participation in the Digital Age: Coping with Information, Participation, and Collaboration Overload', in *End-User Development*, Springer, 2015, pp. 271–275.
- [7] G. Fischer, 'Understanding, fostering, and supporting cultures of participation', *interactions*, vol. 18, no. 3, pp. 42–53, 2011.
- [8] N. Gillani and R. Eynon, 'Communication patterns in massively open online courses', *Internet High. Educ.*, vol. 23, pp. 18–26, 2014.
- [9] J. Reich, D. H. Tingley, J. Leder-Luis, M. E. Roberts, and B. Stewart, 'Computer-Assisted Reading and Discovery for Student Generated Text in Massive Open Online Courses', *J. Learn. Anal.*, vol. 2, no. 1, pp. 156–184, 2014.
- [10] C. G. Brinton, M. Chiang, S. Jain, H. Lam, Z. Liu, and F. M. F. Wong, 'Learning about social learning in MOOCs: From statistical analysis to generative model', *ArXiv13122159 Cs*, Dec. 2013.
- [11] B. De Wever, T. Schellens, M. Valcke, and H. V. Keer, 'Content analysis schemes to analyze transcripts of online asynchronous discussion groups: a review', *Comput Educ*, vol. 46, no. 1, pp. 6–28, 2006.
- [12] D. R. Garrison, T. Anderson, and W. Archer, 'Critical inquiry in a text-based environment: Computer conferencing in higher education', *Internet High. Educ.*, vol. 2, no. 2, pp. 87–105, 1999.
- [13] C. Gunawardena, C. Lowe, and T. Anderson, 'Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing', *J. Educ. Comput. Res.*, vol. 17, no. 4, pp. 395–429, 1997.
- [14] F. Henri, 'Computer conferencing and content analysis', in *Collaborative learning through computer conferencing*, Springer, 1992, pp. 117–136.