# New platform affordances for encouraging social interaction in MOOCs

The "Comment Discovery Tool" interactive visualisation plugin

Philip Tubman
School of Computing and Communications
Lancaster University
Lancaster, UK
p.tubman@lancaster.ac.uk

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

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

***Abstract*—This paper describes the development of a tool which aggregates learner MOOC contributions into a wordcloud. It is designed to serve either as a concept filter or a way to discover new conversations. This is analysed using a measure developed as a heuristic for sociocultural learning in conversations and by a survey. A new pedagogical approach is suggested which adds to the theories behind MOOC pedagogy, by using novel platform affordances to increase active participation.**

***MOOC, Social Learning, DBR, FutureLearn, Visualisation***

## I. Introduction

This paper describes a design-based research (DBR) intervention in the FutureLearn MOOC platform which uses raw computing power in order to visualise all user comments into an interactive wordcloud, such that learners in the MOOC are able to filter concepts, discover conversations and connect with other learners more easily.

Previous research has established that the developing sustained conversations in MOOCs is problematic and that most comments do not receive a reply, which is a basic pre-requisite for social learning [1].

The paper describes the process of design and analysis for a tool which attempts to improve this situation. The results are presented for 259239 conversations and 35 courses, 9 of which are plugin-enabled. There is also a survey which is completed fully by 304 participants in plugin-enabled courses to examine in more detail how the intervention is perceived and how useful it is appraised by the participants.

We conclude with an account of the next steps in the DBR and how this relates to the theoretical framework underpinning the design and analysis.

## II. Theoretical Framework

This project looks at MOOC learning context from the sociocultural viewpoint that there is a wealth of knowledge which is held by the learners themselves, and that each learner will bring something different to the course. The FutureLearn platform differs from other learning environments because the courses are broken down into sequential webpages called 'steps' and importantly there is a place on each step for learners to leave comments and have conversations. This builds on the Pask-Laurillard 'Conversational Framework' [2] where conversations are positioned at the 'Level of Actions' and discussion is directed towards interpreting the immediate learning content.

Learners in MOOCs can be described as self-directed. The ratio of learner to teacher is often 1000:1 so personalised, guided support is almost impossible which makes it important that MOOC platforms are designed so that learners are able to customise their participation to what is most meaningful to them [3].

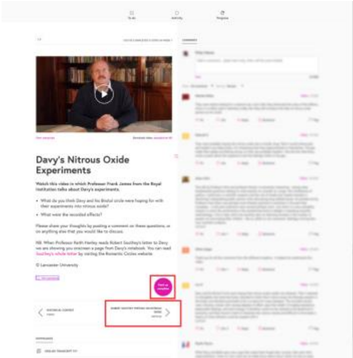


Figure 1. A FutureLearn 'step' with progress buttons highlighted

## III. Comment Discovery Tool (CDT) Plugin

The 'Comment Discovery Tool' (CDT) is an intervention on the FutureLearn platform which visualises the words used in learner comments into an interactive wordcloud, using Python natural language toolkit (nltk) and d3.js.

Learners are presented with a wordcloud of the most frequent 200 words. They are able to click on a word which redraws the cloud based on only comments which include their chosen word. The tool is therefore a concept map, and it is possible to see which words are commonly used together in comments. Learners can choose as many words as they wish in order to filter the whole corpus and choosing 3 words will typically filter a corpus of 10000 comments down to around 10. The tool also provides links to each comment displayed so that learners are able to join the conversation themselves.


Figure 2. Comment Discovery Tool

## IV. METHODOLOGY

DBR is an emerging method in educational research since its inception by the ‘design experiments’ of Ann Brown [4]. In recent years, the idea that learning is situated has become axiomatic for many educational researchers, and DBR is enacted in a real-world context. A firm theoretical grounding is required for any DBR project, such that results can be understood conceptually, and be repeatable.

As described above, the theoretical framework for this project takes a sociocultural perspective of learning, so diversity of participation and sustained interactions are measured. This research project concerns a change of the affordances of the platform, and so a comparison is made between CDT-enhanced courses and those with standard FutureLearn affordances.

### *A. Analysing the Social Dimension*

Conversations in the FutureLearn platform are not threaded. Chua et al. propose a taxonomy of FutureLearn posts based on platform affordances which places them in one of 5 categories (initial, lone, first reply, further reply, initiator reply), where ‘further’ represents sustaining the conversation through turn taking so extending the funds of knowledge [5], [6]. This project extends the taxonomy for posts so it can be used on whole conversational units and all the possible conversation types are displayed below (Table 1). We also add a taxonomy based on unique participants in a conversation to represent diversity [7].

### *B. Analysing the Diversity Dimension*

Unique participants in conversations can be divided into 3 distinct groups, with room to expand upwards. These are conversations with 1 member (“Lone”), conversations with 2 members (“Watercooler”), and conversations with 3-9 members (“Cocktail Party”). These groupings can be used in combination with the ‘social dimension’ as explained above, in order to describe the nature of all the conversations in the MOOC.

TABLE I. POSSIBLE TYPES OF CONVERSATION ON THE FUTURELEARN PLATFORM

| Initial Post (IP) | First Reply (FR) | Further Reply (FurR) | Initiator First Reply (IR) | Initiator Further Reply (IFurR) | Heuristic Type |
|---|---|---|---|---|---|
| ✓ | ✓ | ✓ | ✓ | ✓ | Extended Social |
| ✓ | ✓ | ✓ | ✓ | | Extended Social |
| ✓ | ✓ | ✓ | | | Extended Social |
| ✓ | ✓ | | | | Q&A |
| ✓ | ✓ | | ✓ | | Limited Social |
| ✓ | ✓ | | ✓ | ✓ | Extended Social |
| ✓ | | | ✓ | | Lone |
| ✓ | | | ✓ | ✓ | Lone |
| ✓ | | | | | Lone |

### *C. Length Attribute*

A python script was developed to count length and unique participants in conversational units. These results were analysed using ANOVA and Cohen’s d in order to show statistical significance and impact of these attributes across courses with and without the intervention.

ANOVA and Cohen’s d scores show statistical significance purely based on count data, but the heuristics demonstrate the quality of this significance in relation to the possible conversation types which are afforded by the platform and described by the ‘social dimension’ taxonomy.

### *D. Survey Tool*

In addition to the quantitative measures described above, we surveyed users of the intervention in order to establish whether they perceived the affordances which were designed and what other affordances were perceived, how easily they found the tool to use, whether it encouraged them to comment more on the course and how ‘useful’ they found it. The results of these questions were analysed using a Spearman’s rank order correlation.

## V. RESULTS

### *A. Does the CDT have a statistically significant impact on length and unique participants of conversations?*

257239 conversations were analysed. An ANOVA analysis showed that the unique learners variable was significant, $F(1, 257239)=496.265$, $p=0.00$, and also that the conversation length variable was significant, $F(1, 257239)=601.703$, $p=0.00$. Cohen’s d scores were also calculated for a measurement of impact, and generated a score of 0.15 for unique learners, 0.12 for conversation length. This suggests the CDT has had a small but noticeable impact across the courses in DBR phase 1.

TABLE II. DESCRIPTIVE STATISTICS

| | Courses (n=35) | *N* | *M* | *SD* |
|---|---|---|---|---|
| Unique Learners | no CDT | 225618 | 1.33 | 0.80 |
| | CDT | 31621 | 1.46 | 0.91 |
| Conversation Length | no CDT | 225618 | 1.48 | 1.43 |
| | CDT | 31621 | 1.67 | 1.70 |

### *B. Does the CDT affect the types of conversations on the platform?*

These breakdowns of conversations by type (according to the heuristic measures explained above) demonstrate that courses with the CDT have a larger proportion of the heuristic groupings associated with higher levels of social constructivist learning: extended social conversation, conversations with more members, and fewer lone conversations.

TABLE III. PERCENTAGE OF CONVERSATIONS IN HEURISTIC GROUPINGS

| | | No CDT | CDT |
|---|---|---|---|
| Social dimension | Lone | 78.15 | 71.06 |
| | Q&A | 14.93 | 19.53 |
| | Limited Social | 3.29 | 3.88 |
| | Extended Social | 3.63 | 5.53 |
| Unique participants | Lone | 78.15 | 71.06 |
| | Watercooler | 15.08 | 19.16 |
| | Cocktail Party | 6.75 | 9.76 |
| | Conference | 0.02 | 0.02 |

### *C. Survey Results*

304 people responded fully to a 15-question survey. A Spearman's rank correlation analysis was conducted on the results. Spearman's correlation coefficient is a measure of the strength of monotonic relationships between paired data where 1 would be a perfect positive correlation, 0.6 would be considered strong and 0.4 would be considered moderate [8]. The analysis shows moderate levels of positive correlation between both valuing social interaction and number of times the CDT was used in relation to all the questions relating to learning (0.35 – 0.49). There was a smaller impact in terms of discovery of new people (0.25) and there is no correlation between number of courses done previously. The strongest correlation is between learners who answer positively to the tool helping develop thinking about the material, discovering new conversations through the tool, and the tool encouraging more commenting behaviour (between 0.6-0.7).

## VI. DISCUSSION

The primary finding from the results is that the intervention does have a statistically significant impact across conversation length, conversation type, and unique participants dimension, which is encouraging in this first stage of DBR. However, these statistics cannot guide development by themselves because they do not differentiate between learners who used the tool and those who did not. The MOOC also has a heterogenous population of learners, and some learners do not see the value in learning through social interactions. The survey tool is able to differentiate to some degree in order to establish how the tool helps with deep learning.

The quantitative analysis reveals a moderate positive correlation between learners who see value in social interaction, and the questions about learning qualities, such as developing thinking through discovery and connecting new ideas and conversations. It shows a strong positive correlation between the perception of the designed affordances, and commenting more.

## VII. CONCLUSION

The CDT was designed with a sociocultural theoretical frame, that there is a wealth of knowledge in the participants, and the design used stigmergic principles to harness the emergent swarm intelligence into a coherent 'learning system'. We believe that this focus on platform affordances as the site of self-directed study adds something to the pedagogy of scale, which is missed by post-hoc analyses of learning analytics or clickstream data. The next step is to integrate these affordances in a wider pedagogical approach such that MOOC learning activities can actually be designed around harnessing the micro-contributions of the learning community, and crowdsourcing knowledge.

An interesting development in MOOC pedagogy could encourage learner behaviours that develop an emergent folksonomy of terms specific to the course in question through use of hashtags which would generate a smaller and more thematic corpus.

## REFERENCES


[1] P. Tubman, M. Oztok, and P. Benachour, 'Being social or social learning: A sociocultural analysis of the FutureLearn MOOC platform', presented at the The 16th IEEE International Conference on Advanced Learning Technologies, Austin, Texas, 2016.

[2] D. Laurillard, *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies*. London: New York: Routledge, 1993.

[3] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, 1991.

[4] A. L. Brown, 'Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings', *J. Learn. Sci.*, vol. 2, no. 2, pp. 141–178, 1992.

[5] S.-M. Chua, C. Tagg, M. Sharples, and B. Rienties, 'Discussion Analytics: Identifying Conversations and Social Learners in FutureLearn MOOCs', *MOOC Anal. Live Dashboards Post-Hoc Anal. Long-Term Eff.*, pp. 36–62, 2017.

[6] S. Herring, 'Interactional coherence in CMC', *J. Comput.-Mediat. Commun.*, vol. 4, no. 4, pp. 0–0, 1999.

[7] P. Tubman, P. Benachour, and M. Oztok, 'Design based research approaches towards enhancing social learning practices in MOOC platforms', presented at the 13th International Conference of the Learning Sciences, London, 2018.

[8] I. Weir, 'Spearman's correlation'. [Online]. Available: http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf. [Accessed: 15-Dec-2018].