

Contextualizing the co-creation of artefacts within the nested social structure of a collaborative MOOC

Stian Håkklev¹, Kshitij Sharma^{1,2}, Jim Slotta³, Pierre Dillenbourg¹

1. École Polytechnique Fédérale de Lausanne, Switzerland

2. Faculty of Business and Economics, University of Lausanne, Switzerland

3. Boston College, USA

Abstract. MOOCs have traditionally been seen as providing an individual learning experience, however there is an increasing trend towards enabling social learning in MOOCs. To make online learning at scale more social and collaborative, some MOOCs have introduced cohorts. The interaction between a smaller number of learners, within a cohort, facilitates a richer exchange of experiences and ideas as compared to the effect of “drinking from the fire hose” felt in MOOCs without cohorts. Traditionally, these cohorts have been formed randomly. In this paper, we examine the MOOC “Inquiry and Technology for Teachers”, where we formed cohorts based on student demographics relevant to our course design. Furthermore, these cohorts (which we called Special Interest Groups, SIGs) contained a nested social structure of small teams that worked together on co-creating a final artifact. The different social planes (whole course, SIGs, teams, and individuals) were linked together by pedagogical scripts that orchestrated the movement of ideas and artifacts vertically and horizontally. In this contribution, we analyzed the interaction between these social planes to contextualize the co-creation of artefacts.

Keywords: inquiry-based learning, orchestration at scale, MOOCs, Massive Open Online Courses, group formation, learning analytics, multi-level analysis, scripting collaboration, CSCL

1 Introduction

Massive Open Online Course (MOOC) platforms like Coursera and EdX have gradually added certain social features to their courses, such as peer reviews, discussion forums, and cohorts. Cohorts are course sub-communities, implemented by partitioning forum threads such that participants in a given cohort are only able to see thread replies by other members of the same cohort, to support more intimate and less overwhelming discussions. These cohorts are typically formed by random assignment.¹

In [9], we have described in detail the design of a MOOC for teacher education, in which we sought to create a collaborative knowledge community, where teachers would be able to connect with relevant peers and share professional resources. While inspired by the connectivist MOOCs, we were simultaneously concerned about providing

¹ A separate use case for cohorts, not discussed here, is to present different content to different populations, such as on-campus learners and informal learners.

enough support and scaffolding to lead the students to meet specific learning goals, and not get confused in a too open-ended learning environment. To support a diverse learner population, we designed semantically meaningful cohorts (called Special Interest Groups, SIGs), such as "Secondary Science", or "Elementary English, History, and Social Studies".

These SIGs formed disciplinary sub-communities that supported participants in reflecting on and applying general course theories to their own specific contexts. The design of the course relied on an integration between EdX functionality, and external LTI² components, to enable students to benefit from their larger disciplinary community, while engaging in the co-creation of lesson design documents in small teams. This was formalized through collaborative scripts that described the flow of ideas between different levels, both explicit and implicit.

The initial bootstrapping of lesson design groups was informed by SIG-level brainstorming around relevant resources, and commenting upon lesson design documents from previous courses. The in-progress lesson designs were regularly circulated out to the broader SIG community for constructive peer-review, with prompts informed by the weekly themes. In addition to these formal interdependencies, work by participants in their small design groups was also naturally informed by their own participation in forum discussions and other collaborative knowledge building activities in the broader SIG.

A clearly explicated learning design calls for a targeted approach to learning analytics. In this paper, we will use learning analytics approaches to investigate the extent to which these various groups could make the MOOC their own, and benefit from appropriate learning trajectories and a personally relevant community experience. We will contextualize the co-creation of course artefacts within a multi-level social context, analyzing the impact of the SIGs, the design groups, and individual behaviour on lesson design quality, as indicated by a coding scheme.

In this contribution, we present the design of a MOOC with a nested social structure (Section 3). Furthermore, we present a new coding scheme to assess the quality of the collaborative artifact generated by the MOOC participants, and an analysis framework to analyse the relationships among the different social granularities in the nested structure (Section 4). Finally, we present the relationships among various social levels based on our multi-level analysis framework (Section 5).

2 Related Work

Cohorts in Computer Supported Collaborative Learning (CSCL): The idea of scripting collaboration in forums/asynchronous chats/discussion groups has been studied in detail over many years in CSCL. One central idea to manage a large number of students, is to scaffold the collaborative learning processes using cohorts [5]. The cohorts can be based upon many factors, such as: roles [22], tasks [4, 18], learning context [15], or learners' experience [24]. In the present contribution, we propose a design based on

² Learning Tools Interoperability, a protocol for embedding components in a Learning Management System

semantically meaningful cohorts, based on the contextual (learning) interests of the participants. In the MOOC, “Inquiry and Technology for Teachers”, the SIGs were created based on the disciplines the participants used to teach in their respective institutions.

Online activities as a measure of student behaviour: Previous research has shown that there is a close relation between students’ online behaviour and their success in MOOCs. We list a few examples here. El Badrway et. al., [6] used collaborative multi-regression with online activities (form views, comments, posts, videos watched, quizzes answered) to predict students’ grade. Similarly, Pardo and colleagues[19] used weekly data from on-line activities , such as, number of play/pause events, number of quizzes submitted number of exercises answered correctly/incorrectly to predict students’ performance. Kennedy et.al., [11] used the success rate in the previous assignments to predict students’ success in the next assignments. Coffrin et. al.,[3] used on-line access routines (videos and assessment submission), to predict students’ success. Ren et.al.,[21] used number of sessions, average session length, number log in, number of quiz, number of videos, pauses, total view time, homework problems (time, sessions) to predict students’ performance. Sharma et.al, [23] showed that delay in video watching, assignment and quiz submission, correlated negatively to grade.

Social Network and/or forum text mining: Another stream of research has used the Social Network Analysis (SNA) based methods to predict final grades of students. For example Brown et.al,[1] found that certain cliques in SNA perform better than the others. Other SNA based variables such as betweenness, upvotes, centrality, degree [2, 10]; density, centrality, efficiency, content richness [20]; forum questions, answers, reads, contributions,[25]; forum text [16]; were found to be correlated with students’ performance. Khan and colleagues [12] showed that there is a correlation between students’ grade and their forum access routines. Some researchers have found the forum or Learning Management Systems (LMS) access routines are predictive of students’ grades. For example, reading forums[8], number of forum/LMS visits and time spent on LMS [17, 7, 13], were found to be good predictors of students’ performance in a MOOC.

In this paper, we combine the online activities and the SNA based variables to quantify the participant activities at different levels of social scales of the MOOC (Section 4). We also show how scaffolding the co-creation of the collaborative artifact, and the peer-review process helps the participants produce high quality artifacts (Section 5).

3 MOOC on Inquiry and Technology for Teachers

The MOOC featured in this study was designed to support in-service teachers in their efforts to integrate inquiry and technology into their lessons. Although open to anyone, it was explicitly marketed to in-service teachers, and was designed to build upon their professional experience and respond to their real challenges, providing tools, examples and approaches that could be directly applied within their professional settings.

The course came out of a collaboration between the University of Toronto Schools (UTS), a university-affiliated private secondary school, and the Encore research group led by Dr. Jim Slotta, enabling us to provide an integration of academic and theoretical ideas, with applied practice. Both the UTS principal, as well as a number of their

experienced teachers, contributed to the design and the contents of the MOOC. In particular, we wanted to showcase the specific cases of technology-enhanced inquiry that were happening at UTS, including some that were the result of research collaborations with Dr. Slotta's group, and some that had occurred spontaneously within the school.

3.1 Course design

The following weekly themes were chosen to bring participants into contact with a variety of technologies and pedagogical topics relevant to their teaching: (1) Inquiry and student-centred pedagogy; (2) Designing inquiry activities and assessments; (3) Collaborative learning; (4) Handheld/mobile devices; (5) Knowledge co-construction and student-contributed content; and (6) Inquiry enactment.

The participants engaged with these weekly topics in a number of ways, on different social planes, as depicted in Figure 1. Each week began with a selection of videos, ranging from theoretical and academic to applied and practical. Participants were asked to write an individual reflection on the topic, to connect the theory with their own experience. This was followed by a forum discussion within the SIG, shaped by prompts appropriate to each week's topic, and an individual evaluation of participants' own discussion forum activity.

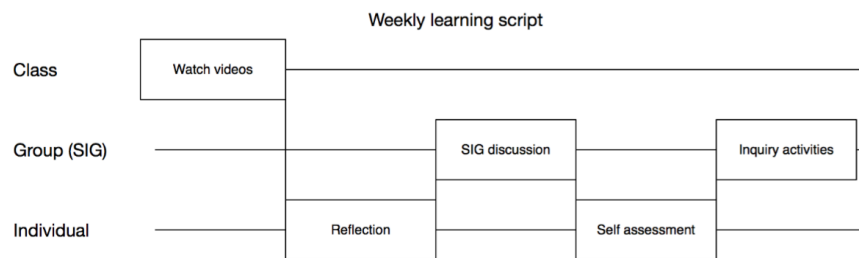


Fig. 1. Pedagogical graph of weekly activities.

Finally, there were a set of inquiry activities each week, which began by contributing to the SIG knowledge base, by for example commenting on old lesson designs, crowd-sourcing relevant technology resources, or brainstorming ideas through voting and commenting. The second part of the inquiry activities centered around the lesson design groups.

3.2 Design groups

The organizing principle of the course was the co-creation of a lesson design, utilizing principles and resources from the course, but adapted to the teacher's own interests and needs. Participants suggested lesson topics, and formed small teams of 2-6 members during the first week of the course. We developed a "collaborative workbench" tool to

The screenshot displays the Collaborative Workbench interface, which is a unified workspace for a MOOC group. The interface is divided into several sections:

- Top Header:** "COLLABORATIVE WORKBENCH" and "UNIVERSITY OF TORONTO".
- Navigation Bar:** "WELCOME", "CONSTRUCTIVE FEEDBACK", "ETHERPAD", "WIKI", and a "LEAVE GROUP" button.
- Left Panel:** A chat window titled "Send email" with a "Chat" button. It shows a list of participants (HUGO, APSA, HANI SAIED, NANDITA) and a message from STIAN: "Hi Hugo, nice to hear from you, and I appreciate your comment".
- Central Workspace:** A Confluence page titled "29: Reaching the research" created by Cory von Encore on Jul 15, 2015. It lists "Team Members:" and contains two numbered prompts:
 - Describe a typical classroom where this lesson might be enacted. What age are the students? What demographic? What kind of diversity is present in the classroom? Are there any particular challenges confronting your students, school or community? --- add your response here (delete this line of text) ---
 - Describe the major theme of the lesson.
- Right Panel:** A sidebar with navigation options: "WELCOME", "CONSTRUCTIVE FEEDBACK", "ETHERPAD", and "WIKI". It also includes a "LEAVE GROUP" button.

Fig. 2. The collaborative workbench, a unified interface to multiple components.

support group collaboration and coordination, which contained all the information and tools needed by the small groups (see Figure 2).

Each week, groups received new prompts and suggestions relevant to the weekly theme, gradually increasing in sophistication. Some sections of the collaborative workbench were private to the group, such as chat, messaging, and scratchpad, while others were bringing in relevant resources from the broader community, such as constructive peer review, and brainstormed resources. Finally, the wiki page, where the group drafted the lesson design, was a public resource, available to the rest of the SIG for review.

The lesson design document was built around a template with the prompts listed below. These were not all made available to the students in the first week, but rather incrementally added to the document.

1. Describe a typical classroom where this lesson might be enacted.
2. Describe the major theme of the lesson.
3. What are the learning goals of the technology-enhanced lesson?
4. Aspects of the design: Student-Centered Design / Peer Collaboration / Use of Handheld or Mobile Computers / Supporting Equity and Diversity
5. What is the activity structure of the lesson?
6. Assessment notes.
7. Enactment notes, and ethics or enactment concerns.

Most weeks, the inquiry activities included peer reviewing lesson design documents in progress. Participants were not asked to rate the quality of unfinished products, but rather to suggest ways of improving the documents, informed by the weekly theme. Given that only a subset of students were actively engaged in the authoring of lesson designs, each lesson design group would receive a large number of aggregated suggestions to inform their regular work.

Figure 3 shows a systematic depiction of the various pedagogical scripts in the MOOC, including the flow of artefacts from previous iterations of the course, and to future iterations, the way activities contribute to the community knowledge base, the initial "bootstrapping" of lesson design teams by community crowdsourcing, and the regular cycling of in-progression design documents through the SIG community for review, and back to the design groups. The script was intentionally developed to account for the fact that one group of participants would want to engage deeply in sustained co-creation, and another group would want a more "traditional" MOOC experience, but that these two groups could not only both be catered to, but also made to be positively interdependent on each other.

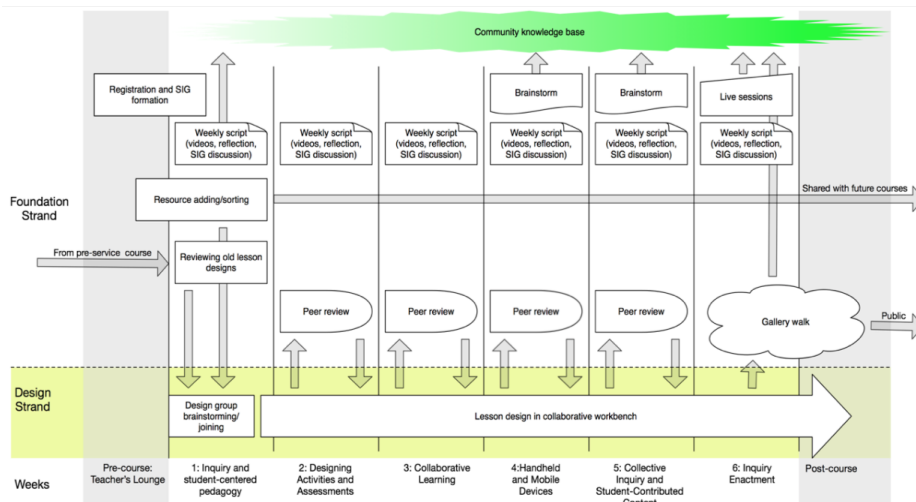


Fig. 3. Pedagogical graph of weekly activities.

4 Variables

Figure 4 shows the nested social structure of the MOOC and the variables we measure at different levels. The top-most level is the whole MOOC learners' population. The second level is the cohorts or, as we will refer to them in the rest of this paper, "Special Interest Groups (SIGs)". The third level contains the design groups. Finally, the fourth level is comprised of the individual learners.

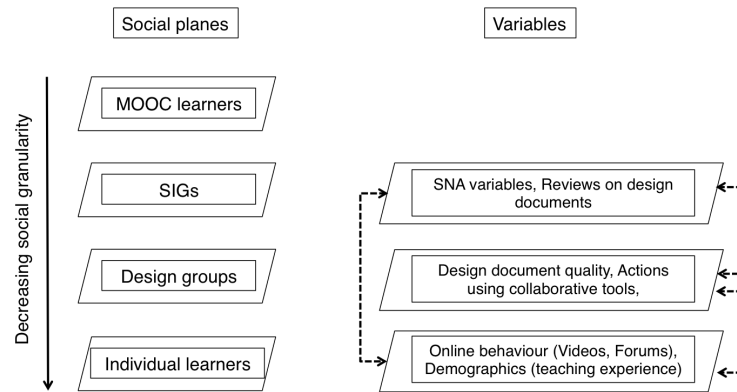


Fig. 4. Social planes (left) and the variables defined at different planes (right). **Left:** the number of members in a social plane decreases as we move from the MOOC participants at the top to the individual learners at the bottom. **Right:** we define variables at three levels, the dashed arrows represent the different relations among the variables computed at the three different social planes.

Each social plane has a set of variables that are either collected at a specific plane or represents the flow of the information between the two levels. The different variables according to their respective planes or the interaction between two planes are as follows: **SIGs in terms of teaching levels (SIG level):** We categorised the SIGs into three categories based on the levels of the education they addressed: K1-6, K7-12, and HighED.

Social Network Variables (SIG level): in the MOOC's forum, each SIG had its own social network. We computed the Social Network Analysis (SNA) variables for each SIG: in-degree, out-degree, and network centrality.

Design document quality (design group level): the final artifact produced by the learners in each design group was the design document, where the members provided details of a course that they were teaching in their respective institutions. Two of the authors coded the quality of these design document (inter-rater reliability = 0.82) based on the following quality metrics, which were derived from the learning objectives of the MOOC. Each criteria was scored from 1-5, with 0 indicating total absence.

- *Learning Objective (LO)*: Level of detail put in the learning objectives mentioned.
- *Activity Design (AD)*: Richness in the design of the activities according to the learning objectives.
- *Coherence (CO)*: Level of coherence in the various parts of the design document.
- *Innovative use of technology (DT)*: Depth of thought put into the innovative use of technology in the design document.
- *Incorporating inquiry-based learning (IB)*: The use of inquiry based learning principles in the design document.

Reviews on the design documents (interaction between SIG and design group levels): in different weeks of the course, the SIG members were given a set a questions to comment on different design documents. For each week that the reviews were requested,

the *reviewMetric* measures how many questions the reviewers answered, and in what detail.

Collaborative actions (design group level): while collaborating on the design documents, the design group members could use various collaborative tools; for example group wiki page, chat tool, and group Etherpad (for collaborative note taking). we measured the amount of activity on these tools per design group.

Video watching behaviour (individual level): there were four different types of videos each week.

- Fireside: Informal weekly introduction, recorded as the course progressed, and reflecting community evolution and emerging questions.
- Academic: Theoretical and conceptual introduction to the weekly theme by Dr. Slotta.
- Principal: Introduction to the weekly theme by principal or vice-principal of a middle school (UTS).
- Practitioner: Interviews with teachers and mini-documentaries from classroom exercises implementing ideas from the weekly theme.

For each video type and each week, we computed: the number of new and old videos watched and the time difference between two video viewing. Besides this, we also counted the different video watching actions from the click stream data: play, pause, seek-back, seek-forward, and speed-change.

Forum access behaviour (individual level): besides computing the SNA variables, we also computed the individual forum actions in terms of the viewing, posting, commenting, and searching behaviour for the whole course.

5 Results and Discussion

In the previous section, we presented 6 different sets of variables based on the activities and the social planes. In this section, we report the relations we found among these variables based on the social plans or the interaction between the two social planes.

5.1 SIG-individual

Video Watching Behaviour for different SIG types We divided the video watching behaviour into two level: (1) the number of videos watched in the same week as they were released; (2) the number of videos watched in the later weeks as they were released. For the number of videos watched in the same week as they were released, overall, members in HighEd SIGs watch the most Fireside ($F[2,4819] = 6.80, p=.001$) and Academic ($F[2,4819] = 10.10, p<.0001$) videos. Whereas, members in K7-12 SIGs watch the least number of Fireside and Academic videos. On the other hand, members in K7-12 SIGs watch the most number of Principal ($F[2,4819] = 5.13, p=.005$) and those in K1-6 the least. Members in the K1-6 SIGs watch the most Practitioner ($F[2,4819] = 10.50, p<.0001$) videos; where as, members of HighEd SIGs watch the least number of Practitioner videos. For the number of videos watched in the later weeks than they were released, overall, members in the HighEd SIGs watch the most Academic ($F[2,4850]$

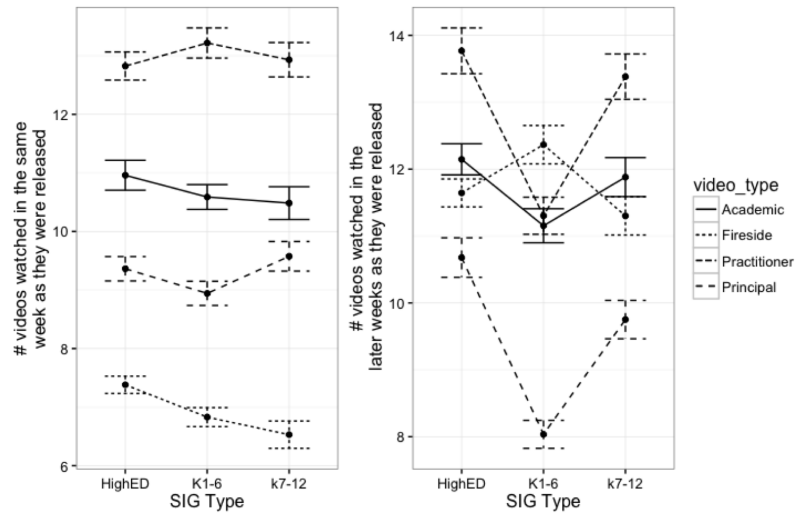


Fig. 5. Left: Number of videos watched in the same week as they were released by the members of different SIG types. **Right:** Number of videos watched in the week later than the week they were released by the members of different SIG types.

= 8.95, $p=.006$), Principal ($F[2,3979] = 14.69$, $p<.0001$), and Practitioner ($F[2,4291] = 7.00$, $p=.0009$) videos; and the members of the K1-6 watch the least number of Academic, Principal, and Practitioner videos. While the members of the K1-6 SIGs watch the most Fireside ($F[2,4819] = 5.07$, $p=.006$) videos, in this case those in K7-12 SIGs watch the least number of Fireside videos. (See Figure 5).

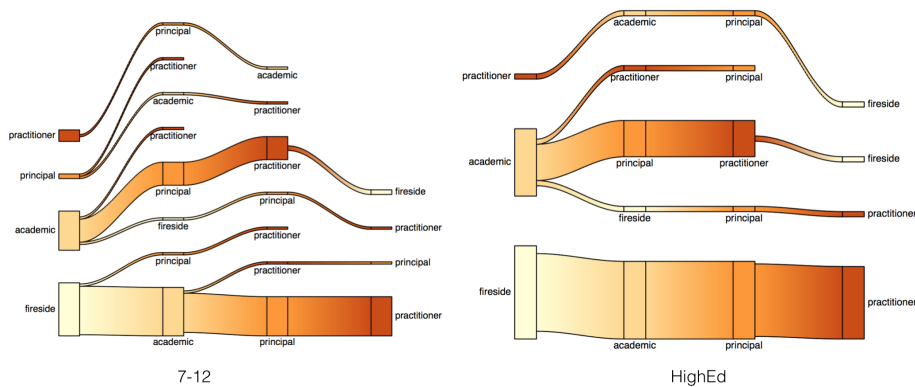


Fig. 6. Sequence of first watches of each video (width indicates number of students).

Student flow through videos Most participants followed the default ordering of the videos in the EdX interface, but to see how these pathways could subtly differ between differently categorized SIGs, we chose a SIG, Secondary Sciences, whose members are representative of the 7-12 mean (SIGs where most members taught between 7th to 12th grade), and another SIG representative of the HighEd population (Foreign Languages and English as a Second Language). Plotting the pathways of participants through the videos (only first views of each videos are taken into account) for week 4, as an example, we can see that HighEd participants are more likely to follow the standard pathway, and focus attention on the academic video. Whereas the 7-12 participants show a much larger diversity of paths, and more focus on school-based videos (practitioner and principal) (see Figure 6).

5.2 SIG-design group

Reviews and design document quality Table 1 shows the relations between the *reviewMetric* and the design document quality. We observe that the *reviewMetrics* from weeks 2 and 3 are correlated to all the quality codes. However, we observe no such relation for the other two weeks. One plausible explanation for this could be the fact that the review questions from weeks 2 and 3 were about generating/brainstorming the ideas and incorporating collaboration in the student activities. These were higher level concepts, which could have been more significant for high ratings in the quality indicators. On the other hand, the review questions from week 4 were about specific topics, for example, using smart phones and tablets in the activity design. This might not had any effect on the design document, as many learners did not plan their activity around such devices. Finally, the review questions from week 5 were about incorporation of inquiry based learning in the lesson plans. This, we hypothesize, was too late in the course time-line to have any effect on the design document quality, as after receiving an insightful feedback from week 5, the design groups might have had to change the whole lesson design to incorporate the new ideas.

SNA and design document quality We observe a significant correlation between the design document quality and the network centrality from weeks 1 and 5 (Table 1). This might be due to the fact that these two weeks correspond to the initial and final weeks of the collaborative work on the design document. High network centrality depicts the fact that all the learners in a SIG were equally contributing to the forums. This could entail the brainstorming conversations among the peers.

Discussion - SIG-design group The two aforementioned results indicate that the scripting of the course might have an effect on the activities of the learners, as well as the interactions between the different social planes. The review questionnaires scripted to be abstract in the beginning of the course, gradually become more concrete. The relationships we found reflect this process. The first two review weeks (1 and 3) had an impact on the quality of the final artifact and the last two did not. Moreover, the script of the course from weeks 1 (initiating collaboration) to 5 (finalising the collaborative artifact) has corresponding actions in the forums as well.

Table 1. Correlations between variables and quality metrics. All the correlations are have at least p-value < 0.1. In the following the significance level is as follows: * < 0.05, ** < 0.01, *** < 0.001

Ind.	Social level	Variable	Estimate	Std.err.
LO	Individual (video)	Seek.Back	-0.007288**	0.002670
		newVideos	0.125865***	0.030932
	Individual (forum)	Forum.Load	-0.020290***	0.008753
		Forum.Search	0.029055	0.016544
	Design group (collaboration)	Collab.Chat	0.081835*	0.034158
		Collab.WikiEdit	0.928531***	0.180442
	SIG (SNA)	Centrality week 1	2.930e+02**	1.098e+02
		Out degree week 1	-1.609e+02**	5.829e+01
		Centrality week 5	5.046e-04*	2.219e-04
	SIG (reviews)	Review Week 1	0.0011257**	0.0003933
Review Week 3		0.0009519**	0.0004802	
AD	Individual (video)	Seek.Back	-0.006289**	0.002604
		newVideos	0.083227**	0.030372
	Individual (forum)	Forum.Load	-0.017708*	0.008426
		Forum.Search	0.033812*	0.015926
	Design group (collaboration)	Collab.Chat	0.0914995**	0.0331748
		Collab.WikiEdit	0.7468573***	0.1752467
	SIG (SNA)	Centrality week 1	2.381e+02*	1.116e+02
		Out degree week 1	-1.331e+02*	5.924e+01
		Centrality week 5	4.746e-04*	2.256e-04
	SIG (reviews)	Review Week 1	0.0010307**	0.0003605
Review Week 3		0.0009525*	0.0004401	
CO	Individual (video)	Seek.Back	-0.006224	0.002739*
		newVideos	0.089395**	0.031171
	Individual (forum)	Forum.Load	-0.020801**	0.008726
		Forum.Search	0.031160	0.016494
	Design group (collaboration)	Collab.Chat	0.087815**	0.035658
		Collab.WikiEdit	0.736609***	0.188362
	SIG (SNA)	Centrality week 1	2.685e+02*	1.159e+02
		Out degree week 1	-1.479e+02**	6.155e+01
		Centrality week 5	5.105e-04*	2.343e-04
	SIG (reviews)	Review Week 1	0.0009799**	0.0003760
Review Week 3		0.0011245**	0.0004591	
DT	Individual (video)	Seek.Back	-0.0055070*	0.0024674
		newVideos	0.075931**	0.028272
	Individual (forum)	Forum.Load	-1.770e-02*	7.781e-03
		Forum.Search	3.490e-02*	1.471e-02
	Design group (collaboration)	Collab.Chat	0.0843472**	0.0306641
		Collab.WikiEdit	0.6282267***	0.1619840
	SIG (SNA)	Centrality week 1	2.308e+02*	1.031e+02
		Out degree week 1	-1.268e+02*	5.474e+01
		Centrality week 5	5.424e-04**	2.084e-04
	SIG (reviews)	Review Week 1	0.0010294**	0.0003357
Review Week 3		0.0010323**	0.0004098	
IB	Individual (video)	Seek.Back	-0.0049693*	0.0024701
		newVideos	0.066964*	0.028568
	Individual (forum)	Forum.Load	-0.013087	0.007833
		Forum.Search	0.035876*	0.014805
	Design group (collaboration)	Collab.Chat	0.0828273**	0.0295544
		Collab.WikiEdit	0.6955376***	0.1561218
	SIG (SNA)	Centrality week 1	2.248e+02*	1.030e+02
		Out degree week 1	-1.235e+02*	5.472e+01
		Centrality week 5	5.496e-04**	2.083e-04
	SIG (reviews)	Review Week 1	0.0009621**	0.0003382
Review Week 3		0.0008912*	0.0004129	

5.3 Within design group

Collaborative tools and design document quality There were many collaborative tools provided to the design groups to facilitate the group work. For example, collaborative etherpads, wikis, and chat-tools. We observed a significant correlation between both the number of wiki-edits, and chat events, with the design document quality. This is not surprising for us, as wiki-edits and chat-activities depict the offline and online sharing and discussing of ideas, respectively. This relation can also be attributed to the design of the course and availability of the collaborative/cooperative tools (Table 1).

5.4 Design group - individual

Video watching behaviour and design document quality We observe a significant negative correlation between the number of seeking-back events on the videos and the design document quality (Table 1). This might have stemmed from the fact that the seeking-back behaviour is indicative of higher perceived difficulty [14]. Those groups who had difficulties in understanding the content also had lower design document quality. Moreover, there was a significant positive correlation between the number of new videos watched and the design document quality. Those groups who watched more videos also had higher design document quality. These relations show the contribution of having a mutual-understanding, achieved by watching videos, of the video material while co-creating the artifact.

Forum behaviour and design document quality We observed a positive correlation between number of forum searches and the design document quality. However number of forum visits were negatively correlated to the design document quality. These two relations indicate that only visiting/reading/contributing to the meaningful threads was helpful in co-creation of the design document.

6 Conclusions

In this paper we have described an innovative MOOC design, with novel technologies and pedagogical scripts that allowed participants with similar disciplinary interests to find each other, and which supported both intensive small-group co-creation, while at the same time letting participants benefit from a larger community of peers (Section 3). We introduced a new qualitative coding scheme to assess the co-created design documents produced by the different design groups. Finally we have introduced a framework for multilevel analysis, where the design document quality is considered as a dependent variable, and we have used various process variables from different social planes of the course to explain the relationship among these social planes as well as the different design document quality levels (Section 4).

Having semantically meaningful SIGs in the course had two effects: 1) on participants' actions, and 2) on the design document quality. We provided evidence for these two effects in two different ways: 1) by showing the differences in actions of SIG members, and 2) by showing the relationship between actions of design group members, and the quality of their design document (Section 5).

One of the interesting observations we found was that some of the review prompts did not show the strong positive correlation with high quality design documents that we had expected. We observed that the prompts with high levels of abstraction (brainstorming and collaboration) were positively correlated with design document quality while prompts related to specific technologies and pedagogies (mobile devices and student generated content) were not correlated with the quality. One plausible explanation for the latter could be the timing of the prompts. The specific prompts were towards the end of the MOOC, when the design groups were too advanced in their documents to be able to incorporate new ideas. We plan to investigate the feedback uptake by the design groups in the future.

There is a growing interest towards MOOCs with complex social structures, where participants benefit from small group collaboration, as well as larger scale communities of interest. This contribution presents one such example and shows the contextualisation of data within the nested social structure. The authors hope that this contribution exemplifies forthcoming MOOCs with innovative social and pedagogical scenarios. In conclusion, this multi-level analysis has opened a few new directions for further investigations and interventions. For example, review uptake as mentioned above, as well as focusing on individual learning gains and small group collaborative mechanisms.

References

1. BROWN, R., LYNCH, C. F., EAGLE, M., ALBERT, J., BARNES, T., BAKER, R., BERGNER, Y., AND McNAMARA, D. Good communities and bad communities: Does membership affect performance. In *Proceedings of the 8th International Conference on Educational Data Mining (2015)*, pp. 612–614.
2. BYDŽOVSKÁ, H. Towards freshmen performance prediction. In *Proceedings of the 8th International Conference on Educational Data Mining (2015)*.
3. COFFRIN, C., CORRIN, L., DE BARBA, P., AND KENNEDY, G. Visualizing patterns of student engagement and performance in MOOCs. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge (2014)*, ACM, pp. 83–92.
4. CRAMPORN, C. An evaluation of formal and underlying factors influencing student participation within e-learning web discussion forums. In *Proceedings of the Fourth International Conference on Networked Learning (2004)*, pp. 417–423.
5. DILLENBOURG, P. Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In *Three Worlds of CSCL: Can We Support CSCL?*, P. A. Kirschner, Ed. Heerlen: Open Universiteit Nederland, 2002.
6. ELBADRAWY, A., STUDHAM, R. S., AND KARYPIS, G. Collaborative multi-regression models for predicting students' performance in course activities. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (2015)*, ACM, pp. 103–107.
7. GARCÍA-SAIZ, D., AND ZORRILLA, M. A promising classification method for predicting distance students' performance. In *Proceedings of the 5th International Conference on Educational Data Mining (2012)*.
8. GUNNARSSON, B. L., AND ALTERMAN, R. Predicting failure: a case study in co-blogging. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (2012)*, ACM, pp. 263–266.
9. HÅKLEV, S., AND SLOTTA, J. D. A principled approach to the design of collaborative MOOC curricula. In *Digital Education: Out to the World and Back to the Campus. EMOOCs 2017. (2017)*, pp. 58–67.

10. JOKSIMOVIĆ, S., MANATAKI, A., GAŠEVIĆ, D., DAWSON, S., KOVANOVIĆ, V., AND DE KERKEI, I. F. Translating network position into performance: importance of centrality in different network configurations. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (2016), ACM, pp. 314–323.
11. KENNEDY, G., COFFRIN, C., DE BARBA, P., AND CORRIN, L. Predicting success: how learners' prior knowledge, skills and activities predict mooc performance. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (2015), ACM, pp. 136–140.
12. KHAN, T. M., CLEAR, F., AND SAJADI, S. S. The relationship between educational performance and online access routines: analysis of students' access to an online discussion forum. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (2012), ACM, pp. 226–229.
13. KIM, J. H., SEODAEMUN-GU, S., PARK, Y., SONG, J., AND JO, I.-H. Predicting students' learning performance by using online behavior patterns in blended learning environments: Comparison of two cases on linear and non-linear model. In *Proceedings of the 7th International Conference on Educational Data Mining* (2014).
14. LI, N., KIDZINSKI, L., JERMANN, P., AND DILLENBOURG, P. Characterising MOOC video behaviours with video interaction styles: What do they tell. In *Proceedings of the 10th European Conference on Technology Enhanced Learning* (2015).
15. LIM, W.-Y., HEDBERG, J. G., YEO, J. A.-C., AND HUNG, D. Fostering communities of practice—a case study of heads of it departments. In *The Annual Convention of the Association for Educational Communications and Technology* (2005).
16. LUO, J., SOROUR, S. E., GODA, K., AND MINE, T. Predicting student grade based on free-style comments using Word2Vec and ANN by considering prediction results obtained in consecutive lessons. In *Proceedings of the 8th International Conference on Educational Data Mining* (2015).
17. McCUAIG, J., AND BALDWIN, J. Identifying successful learners from interaction behaviour. In *Proceedings of the 5th International Conference on Educational Data Mining* (2012).
18. McDougall, M. J., Nason, R. A., AND McROBBIE, C. J. Growth of teacher knowledge: The promise of CSCL. In *AARE 2004 International Education Research Conference, Melbourne, Australia* (2004), Australian Association for Research in Education.
19. PARDO, A., MIRRIAH, N., DAWSON, S., ZHAO, Y., ZHAO, A., AND GAŠEVIĆ, D. Identifying learning strategies associated with active use of video annotation software. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (2015), ACM, pp. 255–259.
20. PAREDES, W. C., AND CHUNG, K. S. K. Modelling learning & performance: a social networks perspective. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (2012), ACM, pp. 34–42.
21. REN, Z., RANGWALA, H., AND JOHRI, A. Predicting performance on MOOC assessments using multi-regression models. In *Proceedings of the 9th International Conference on Educational Data Mining* (2016).
22. SCHELLENS, T., VAN KEER, H., DE WEVER, B., AND VALCKE, M. Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *International Journal of Computer-Supported Collaborative Learning* 2, 2 (2007), 225–246.
23. SHARMA, K., JERMANN, P., AND DILLENBOURG, P. Identifying styles and paths toward success in MOOCs. *International Educational Data Mining Society* (2015).
24. STEPHENS, A. C., AND HARTMANN, C. E. Using an online discussion forum to engage secondary mathematics teachers in teaching with technology. In *American Educational Research Association Annual Meeting* (2002).
25. TOMKINS, S., RAMESH, A., AND GETOOR, L. Predicting post-test performance from online student behavior: A high school MOOC case study. In *Proceedings of the 9th International Conference on Educational Data Mining* (2016).