

# Influential Factors for Managing Virtual Groups in Massive and Variable Scale Courses

Luisa Sanz-Martínez<sup>1</sup>, Yannis Dimitriadis<sup>1</sup>, Alejandra Martínez-Monés<sup>1</sup>, Carlos Alario-Hoyos<sup>2</sup>, Miguel Bote-Lorenzo<sup>1</sup>, Bartolomé Rubia-Avi<sup>1</sup>, Alejandro Ortega-Arranz<sup>1</sup>

<sup>1</sup>Research Group GSIC-EMIC, Universidad de Valladolid, Valladolid, Spain

<sup>2</sup>University Carlos III, Madrid, Spain

{luisa,alex}@gsic.uva.es, {yannis,migbot}@tel.uva.es, amartine@infor.uva.es, calario@it.uc3m.es, brubia@pdg.uva.es

**Abstract**—Integration of collaborative learning in MOOCs (Massive Open Online Courses) is an open research challenge. However, team formation and subsequent management are complex tasks that depend on multiple factors, both pedagogical and technological. To fulfill these tasks, it would be helpful to provide teachers with supporting tools. This paper analyzes the factors influencing the formation of teams in MOOCs that can be taken into consideration in the design of this type of supporting tools. The paper presents a proposed classification and illustrates their need and utility by a scenario.

**Keywords**— Groups; CSCL; Collaborative Learning; Massive Scale; MOOC; Group Formation;

## I. INTRODUCTION

According to some authors, MOOCs (Massive Open Online Courses) have led to a change in the model of higher education [1] and a democratization of access to education [2]. Other authors criticize the low instructional quality of these courses [3], indicating their high dropout rate [4] and identifying important research challenges related to the promotion of social interactions that generate knowledge [5], or the development of new pedagogical approaches that take advantage of the large scale [6].

Since the appearance of the first MOOC in 2008 (Connectivism and Connective Knowledge - CCK08), multiple researchers have tried to include connectivist or constructivist educational theories in these courses. In some cases, the authors' purpose was to exploit the possibilities of social interaction offered by the large scale [7], or improve the quality of the learning experience [8]. In others, the purpose was to overcome the deficiencies of these types of courses, such as the low level of student engagement [9] or the low percentage of students completing the course [10]. However, the results of these experiences [11] show that apart from improvements in the courses' completion rates the overall objectives have not been met [12].

CSCL (Computer-Supported Collaborative Learning) [13]

---

This work has been made possible thanks to the partial support of the Spanish projects TIN2014-53199-C3-2-R y VA277U14.

has been extensively studied in educational environments of small and medium scale. However, due to the inherent features of MOOCs, such as their large and variable scale, their students' heterogeneity and low engagement level [14], the implementation of collaborative learning strategies in this area presents many difficulties [5], and research challenges [6].

In this line, the group formation problem in massive contexts is attracting the interest of researchers who explore several techniques and approaches, with the intention of improving social interactions and the student's engagement level. Zheng et al. [9] use two types of algorithms: random and based on initial surveys of students, to form groups to this end. Sinha [10] proposes the use of social network analysis and machine learning techniques to dynamically form teams. The work from Spoelstra et al. [15] analyzes the formation of groups in project-based learning, considering the background knowledge of students, their preferences and personality. Furthermore, Wen et al. [16] study what are the features that distinguish a successful team of students. This wide range of perspectives suggests that there are many factors that can be considered in the formation of groups. Moreover, due to the low and variable level of students' engagement and their high dropout rate, the group maintenance is difficult, even if the groups are formed using sound baseline criteria. Therefore, a method for the dynamic handling of teams (initial formation and eventual restructuring) could contribute to the solution of the aforementioned problem.

Our interest focus on investigating how to design tools that help MOOC's teachers<sup>1</sup> to form teams and also to restructure these teams, if necessary, during the course enactment. To undertake this overall objective, in the first phase of our project, we propose to analyze the context and the various factors that may be considered for the design of such tools. The purpose of this paper is, firstly, to identify and categorize

---

<sup>1</sup> We use the term *teacher* to refer the different users involved in the creation and management of MOOC, such as instructional designers, tutors, teaching assistants, etc.

these factors mapping them to different levels of abstraction and secondly, to illustrate the usefulness of our proposal contextualizing it in a realistic environment through a scenario. The use of scenarios allows us to identify properties that characterize the MOOCs context, detect the requirements of the tools and promote debate on the type of functionality that should be included in these tools.

## II. PROPOSED CLASSIFICATION

The project is currently in a first exploratory phase carried out by a literature review. For this purpose, we analyzed articles from: (i) a search on Scopus and on Web of Knowledge of combinations of the terms "CSCL", "Group formation", "Teams" and "MOOC" and (ii) references in articles previously selected. The retrieved references have been filtered based on the number of citations and a critical analysis of the abstract to ensure whether it was within our field of interest. The information obtained was analyzed and synthesized within our classification proposal.

For the creation of the proposed classification, an iterative analysis process of the related literature was followed in order to obtain information about: (i) main aspects and perspectives in relation to the formation of groups in CSCL, (ii) characteristics of the MOOCs' environment that can influence the group formation and group restructuring, (iii) challenges related to the integration of CSCL in MOOCs' environments and (iv) prospects from research groups that have addressed the problem of the group formation in MOOCs.

Once analyzed the collected information we identified four different categories or levels of abstraction in which the factors to be considered can be framed. Subsequently, we generated through an iterative process a classification scheme (shown in Fig 1), including two perspectives: (a) a hierarchical classification, and (b) a perspective with several levels of abstraction (from pedagogy to technology).

Fig. 1 depicts a first hierarchical perspective showing two dimensions: the *technological* factors (related to the design and implementation) to be considered when incorporating the rest of the factors in a computational tool; and the *pedagogical* factors that the teacher could use for grouping students to perform collaborative activities. These pedagogical factors can be further classified into different categories, depending on the phase of the course lifecycle. *Learning design* factors are typically considered when the teacher designs the course, although they could be reconsidered during the course enactment. The student data captured at the beginning of the course and whose values are not monitored nor updated during the course are considered *static factors*. Finally, the factors related to the *course activity* are the data that emerge by monitoring the students' progress during the course. The second perspective of Figure 1 sorts the factors according to

different levels of abstraction in pedagogy (course design, course development and student as a subject) and technology.

The proposed classification poses the relevance of the pedagogical factors, since they occupy 18 out of 21 categories. Moreover, those factors related to the dynamic activity of the course are critical and differentiators in the MOOCs' environment and may be a help in the dynamic restructuring of groups. Therefore, we believe these dynamic factors are the most important ones to reach our main goal: the development of supporting tools that can be used by teachers for the dynamic management of teams in MOOCs.

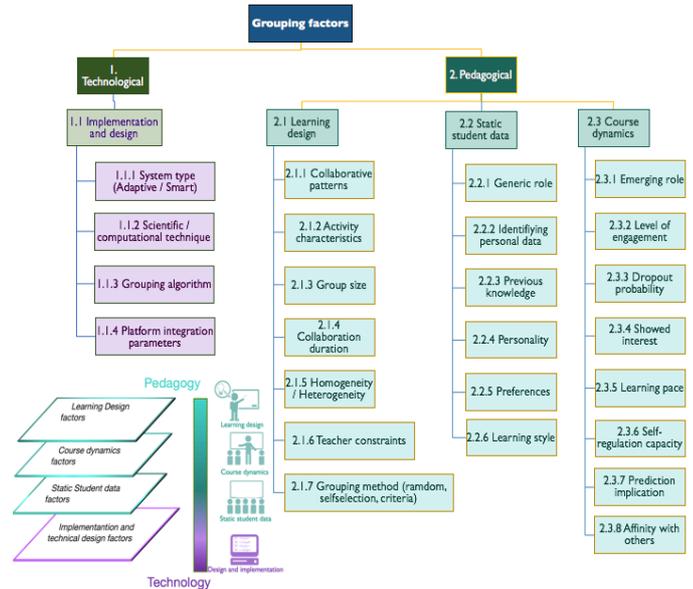


Fig. 1. Classification of influencing factors for managing groups under two perspectives: hierarchical and levels of abstraction.

## III. ILLUSTRATIVE SCENARIO

The following fictional scenario is inspired by data collected from the literature (ratios and typical MOOCs' measurements). Such scenario aims to illustrate how the use of the above identified factors could help in the design of tools for creating and managing virtual teams. The description includes, in parentheses, the factor code (see Fig. 1) to consider in the situation described. It should be noted that the scenario intends not to be exhaustive using all the factors identified, but rather realistic about the description of plausible situations.

University Duero, with experience in classroom education, will introduce new degrees based on online learning through a virtual education platform. To attract students to the online degree, the University decides to launch a series of MOOCs (have not yet decided on which platform). MOOC participants must complete a series of activities to pass the course and perform an identity authentication to finally obtain a verified

certificate. Such students that achieve the verified certificate will get a recognition of official university credits.

Victor, a teacher with previous experience in virtual learning environments, is responsible for the planning and design of the content and activities of a MOOC called "Diet Techniques", which will permit obtain credits within the official online degree "Human Nutrition and Dietetics". The MOOC will last 8 weeks and during the course should be achieved certain competencies of the official course which will be the subject of partial recognition of credits.

The teacher begins the design phase of the course with the clear conviction that it will contain collaborative activities, since he considers collaboration as beneficial to improve the quality of learning. He plans to perform an evaluable activity each week and he is thinking on some learning design patterns (2.1.1) that have been applied in collaborative classes and he would like to apply in the online course. In order to not to be limited to use only peer review, he begins to consider the possibility of including an activity that uses a puzzle or pyramid pattern. Also he wants to include productive activities in small groups (2.1.3), taking into account that such activities must have a maximum duration of one week (2.1.4). Not knowing how many students will enroll, or what will be their behavior during the course, it is difficult to make a preliminary design of activities (2.1.2) and groups to form. That is why, he thinks it would be useful to have some initial information from the students before the start of the course, information relating to their prior knowledge (2.2.3), any personal information such as age, physical location (2.2.2), even some details about preferences (2.2.5) regarding their study schedules, learning style, or the role in which they feel most comfortable when they are working as a team. His intention is to create groups using his criteria (2.1.7) to obtain heterogeneous teams (2.1.5), since, in his view, creating "homogeneously heterogeneous" teams would be the best option for the development of the course. However, if the MOOC has some success, the teacher believes that the collaboration of several teaching assistants will be needed, and the task of creating teams could be very complicated. On the other hand, he is aware that many students enrolled in a MOOC never perform an activity on it and decides that maybe the right thing is to start with an individual activity that allows he to monitor the students' behavior (2.3. 2 and 2.3.1). In such activity, he might include a common forum for discussion and he prepares it so that it allows him to get some more information about common interests (2.3.5), or possible affinities (2.3.4) among students who might fit them in the same team. At this moment he realizes that he needs to know in advance what platform will be used to deploy the design, since the type of resources, activities and means to materialize the groups of students will be conditioned by it (1.1.2).

Finally, he decides to make a design as shown in Table 1, with individual activities that include peer review at all odd weeks (except the first one), a practical assignment in groups of 4-5 students at weeks 2 and 4, an activity using a puzzle pattern in week 6 and an activity using a pyramid pattern in

week 8. As he has doubts whether the planned design can be carried out effectively, he decides to unfold the activities week by week, and be ready to perform interventions he deems necessary so that the exploitation of the course by the participants is appropriate in accordance with the objectives of learning.

TABLE I: TYPES OF PLANNED ACTIVITIES

Activity	Features	
	Individual/Team	Peer-review?
1	Individual	N
2 y 4	Team (productive)	N
3, 5 y 7	Individual	S
6	Team (puzzle)	N
8	Team (pyramid)	N

The MOOC is deployed on a massive platform that records 1400 enrolled students of which 400 (28.5%) filled in the initial survey of personal data and only 150 (10.7%) perform the first solo activity. Of those 150, 15 had not completed the initial survey. With this information and with the help of several colleagues, he performs a manual configuration based on his own criteria in the structure of the first groups and make virtual teams of 4 or 5 students with the limited features offered by the platform. The development of this activity raises numerous complaints from students about the lack of participation by their teammates, and incidences of students who had not been included in any team (not having participated in the previous activities) and wanted to perform this activity. Individual activity 3 is completed by 90 students (6.4%) of which 4 had not made any previous activity.

For the peer review the teacher asks each student to choose 2 tasks of other students to review, but more than 50% of the tasks are without any revision. For this reason, Victor realizes that to handle the information necessary for the formation of groups for activity 4 will be impossible to do manually, since to do so by applying his criteria requires information on the evolution of the activity of the students until that date (2.3) conjugated with information about students (2.2) on which apply his own criteria (2.1.6). This leads him to modify the original design and reconvert all activities in individual, eliminate peer review and without collaboration, due to the lack an automated tool. This tool could have allowed Victor to set up the groups, according to the criteria he wanted, monitor students' activity and reconfigure teams based on the progress of the activity and behavior of the students. When the course ends, he is convinced that not all competencies expected to be

achieved in the MOOC were acquired by the 70 students (5.3%) who passed the course obtaining a verified certificate.

This scenario has shown that teachers who want to include collaborative learning MOOCs need supporting tools that enable them to meet this challenge. These tools can take the form of guidelines or design patterns, and give automatic or semiautomatic support to teachers in managing groups. The tools should allow the teacher to create groups according to various criteria (regarding the design of learning, student characteristics, etc.), taking into account different input parameters. These tools should monitor the activity as the course progress, allowing to incorporate predictive techniques for detecting the level of student engagement, or their probability of dropout; generate warnings, or perform restructurings, when certain levels of degradation are identified, established by the teacher in the composition of the groups. The scenario illustrates that the problem of working with groups of students is not limited to their initial creation, but includes monitoring the team dynamics. And finally, it can also be appreciated that the considerations to take into account for the management of groups belong to different pedagogic categories related to learning design, the dynamics of the course and the profile data of the students.

#### IV. CONCLUSIONS AND FUTURE WORK

MOOCs could benefit from the advantages of collaborative learning if teachers had the tools to manage groups dynamically. The proposed classification has allowed us to delve into the complex problem of creating and maintaining groups in MOOC environments. It has been shown the many factors that can influence such groups and that could be considered to create tools to support teachers in carrying out this task. Factors directly related to the dynamics of the course especially characterize the MOOC environment and can be critical in the design of such tools.

In environments with a fluctuating and massive scale teachers need some type of support, as shown in the illustrative scenario. This support could include automatic or semiautomatic tools to help manage the vast amount of information, which could hardly be handled manually, or patterns or design guidelines created specifically for this type of course. Without this support, it will be difficult that the deployment of collaboration can be effective in such an environment

We plan to continue refining our classification model based on the processing and analysis of semi-structured interviews with experts on CSCL currently working in the design and coordination of MOOC. In addition, further analysis of existing cases helps us triangulate the proposal. As a continuation of our study, we propose an intervention in a MOOC designed by our research group, from which we can draw conclusions to begin the iterative and gradual design of tools to support MOOCs' teachers in managing groups.

#### REFERENCES

- [1] T. Brown, "Exploring new learning paradigms - A reflection on Barber, Donnelley and Rizvi (2013): "An avalanche is coming: Higher education and the revolution ahead".", *International Review of Research in Open and Distributed Learning*, 16 (4), pp. 227-234, 2015.
- [2] A. Balula, "The promotion of digital inclusion through MOOC design and use: a literature review", *Indagatio Didactica*, 7(1), 2015.
- [3] A. Margaryan, M. Bianco and A. Littlejohn, "Instructional quality of Massive Open Online Courses (MOOCs)", *Computers & Education*, 80, pp. 77-83, 2015.
- [4] D. F. O. Onah, J. Sinclair, and R. Bollat, "Dropout rates of Massive Open Online Courses: behavioural patterns" in 6th International Conference on Education and New Learning Technologies, Barcelona, Spain, 7-9 Jul 2014, pp. 14-15.
- [5] K. Manathunga and D. Hernández-Leo, "Has research on collaborative learning technologies addressed massiveness? A literature review." *Educational Technology & Society*, 4522, pp. 1-14, 2015.
- [6] P. Dillenbourg, A. Fox, C. Kirchner, and M. Wirsing, "Massive Open Online Courses: Current state and perspectives." *Dagstuhl Manifestos*, 4(1), pp. 1-27, 2014.
- [7] J. Blom, N. Li, and P. Dillenbourg, "MOOCs are more social than you believe." *eLearning Papers*, 33, May 2013, pp. 1-3.
- [8] G. Conole, "MOOCs as disruptive technologies: strategies for enhancing the learner experience and quality of MOOCs." *RED - Revista de Educación a Distancia*, 39, 2013.
- [9] T. Sinha, "Together we stand, together we fall, together we win: Dynamic team formation in massive open online courses." *The Fifth International Conference on the Applications of Digital Information and Web Technologies* pp. 107-112, 2014.
- [10] Z. Zheng, T. Vogelsang, B. Berlin, and N. Pinkwart, "The impact of small learning group composition on student engagement and success in a MOOC.", in *Proceedings of the 8th International Conference of Educational Data Mining* pp. 500-503, 2015.
- [11] J. Mackness, S.F.J. Mak, and R. Williams, "The ideals and reality of participating in a MOOC." *Learning*, 10, December 2011, pp. 266-274.
- [12] M. Wen, "Investigating virtual teams in Massive Open Online Courses: Deliberation-based virtual team formation, discussion mining and support." PhD Thesis Proposal, Carnegie Mellon University, 2015.
- [13] T. D. Koschmann, "CSCL, theory and practice of an emerging paradigm", Routledge, 1996.
- [14] L. Sanz-Martínez, A. Ortega-Arranz, Y. Dimitriadis, J. A. Muñoz-Cristobal, A. Martínez-Monés, M.L. Bote-Lorenzo and B. Rubia-Avi, "Identifying factors that affect team formation and management in MOOCs", paper accepted at *Intelligent Support for Learning Groups in the International Conference on Intelligent Tutoring Systems*, 2016.
- [15] H. Spoelstra, P. Van Rosmalen, and P. Sloep, "Toward project-based learning and team formation in open learning environments." *Journal of Universal Computer Science*, 20(1), pp. 57-76, 2014
- [16] M. Wen, D. Yang, and C.P. Rose, "Virtual Teams in Massive Open Online Courses.", in *Proceedings of the International Conference on Artificial Intelligence in Education*, Vol. 9112, pp. 820-824, 2015