The Impact of Small Learning Group Composition on Student Engagement and Success in a MOOC

Zhilin Zheng*
Department of Computer Science
Humboldt-Universität zu Berlin
Berlin, Germany
zhilin.zheng@hu-berlin.de

Tim Vogelsang*
iversity GmbH
Bernau bei Berlin, Germany
t.vogelsang@iversity.org

Niels Pinkwart
Department of Computer Science
Humboldt-Universität zu Berlin
Berlin, Germany
niels.pinkwart@hu-berlin.de

ABSTRACT

A commonly known and widely studied problem of massive open online courses (MOOCs) is the high drop-out rate of students. In this paper we propose and analyze the composition of small learning groups as a solution to this problem. In an experiment, we composed such small learning groups in a MOOC context using two methods: Random grouping and grouping by an algorithm that considers selected student criteria. Further, a flipped classroom course was conducted on-campus with a local student group using the MOOC. We compared all three approaches to a control condition using two measures: Drop-out rate and learning performance. The empirical results give an indication, yet no hard evidence, that small groups might reduce student drop-out rates.

Keywords

MOOC; Group Composition; Learning Analytics; Drop-out Rate.

1. INTRODUCTION

MOOC providers, such as Coursera, EdX and iversity, reach course enrolments of up to tens of thousands of students using scalable techniques like lecture videos and quizzes [7]. This massive scale reduces the opportunities for interaction with course instructors. Completion rate, a commonly used (yet debatable) measure of student success, is reported to be less than 13 percent in most MOOCs [3], which has recently attracted extensive studies in order to discover reasons behind this problem [5; 8; 11]. Social connections and collaboration between MOOC students also fall far below expectations. Only 5-10 percent actively participate in course forums [9]. At this point, group formation might help by leading to the creation of informal social ties [4] as well as improving social skills [10].

The composition of small learning groups has already been tested in online learning contexts and local meeting scenarios (i.e. face-to-face groups). In general, self-selected, random and algorithm-based group composition are commonly applied. Algorithm composed groups typically bring together students with either heterogeneous or homogeneous criteria (e.g. based on learning style, personality and demographic information) using technologies such as GT [1] or Swarm Intelligence [2]. Unlike the case with randomly composed or self-selected groups, students' information must be preliminarily collected and then provided to the composition algorithm.

In order to investigate the impact of small learning groups on drop-out rate and learning performance, we conducted a grouping experiment on the iversity.org platform. Specifically, we tested three grouping approaches, all in the same MOOC: 1) automated group composition using an adapted k-means clustering algorithm, accounting for both homogeneous and heterogeneous student criteria; 2) random group composition; and 3) an on-campus flipped classroom approach. This paper describes the results in the three conditions concerning drop-out rates, learning performance and student engagement. The employed algorithm is easy to implement and has low computational costs. In the experiment, we made use of only free and minimal intervention (email) and collaboration methods (email, VoIP, social media). Hence the organizational burden for developers, instructors and students was reduced to a minimum. The experiment is thus scalable and reproducible within many learning environments.

2. METHODOLOGY

2.1 Research Objectives

Empirically, we investigated the following three research questions:

1) Student engagement: Will MOOC students assigned to online groups (without further moderation) be engaged in online collaboration?

2) Drop-out rate: Will random or algorithmic grouping of MOOC students decrease the drop-out rate?

3) Learning performance: Can random or algorithmic grouping lead to higher learning performance, as measured in quizzes and homework scores?

2.2 Experiment Procedure

For conducting the experiment, we chose the second iteration of the course "The Fascination of Crystals and Symmetry", which was offered on the iversity.org platform. This is an introductory course to crystallography held by Dr. Frank Hoffmann (University of Hamburg). Since the course offered open discussion questions, it seemed well suited to engage students in group interaction. It had 3,209 enrollments in total, out of which 771 (i.e. 24.03%) were actively engaged throughout the course.

After the start of the course, 80 percent of the participants received a grouping survey via email asking for information about gender, timezone, language, personality, learning goals (general or in-depth) and their preferred collaboration method (local, email, Facebook, Google+ or Skype). The remaining 20 percent of the course received a motivational survey instead and served as a control condition. One week after the course start, students who provided sufficient answers to the grouping survey were assigned to groups of size 10 by our algorithm and received a second email a few days later. Those who did not respond but had a Facebook account were still randomly assigned to groups. The second email presented the other group members with their personal

* Zhilin Zheng and Tim Vogelsang contributed equally to this work.

descriptions as given in the survey. Further, the email contained a link to the first open discussion question of the course material and a link to their group (if applicable). Students from the control conditions, without or with insufficient grouping survey responses, were not assigned to groups. In addition, the course was held by Dr. Hoffmann as a flipped classroom at the University of Hamburg with approximately 65 students who watched the online lectures at home and met in-class for discussion. Out of these 65 students, 7 used their university account to sign-in to iversity and were anonymously included into our dataset. The other 58 students were not explicitly included. They either used private email addresses or did not sign up to the online course. This (relatively complex, but ecologically valid) assignment procedure of students to seven different conditions is summarized in Figure 1 and Table 1.

Table 1. Student conditions

Condition	Collaboration Method	Description
"Algorithm composed groups" (AlgoCG)	According to preference	Grouping survey, responded sufficiently grouped by algorithm
"Randomly composed groups" (RandCG)	Facebook	Grouping survey, not responded, Facebook user, grouped randomly
"Flipped classroom group" (FlippCG)	Local at University of Hamburg	Attended flipped classroom with the instructor
"No grouping - no answer" (NoG-NA)	none	Grouping survey, not responded, not grouped
"No grouping - insufficient answer" (NoG-IA)	none	Grouping survey, responded insufficiently, not grouped
"No grouping - control group - responsive" (NoG-CG-R)	none	Motivational survey, responded, not grouped
"No grouping - control group - nonresponsive" (NoG-CG-NR)	none	Motivational survey, not responded, not grouped

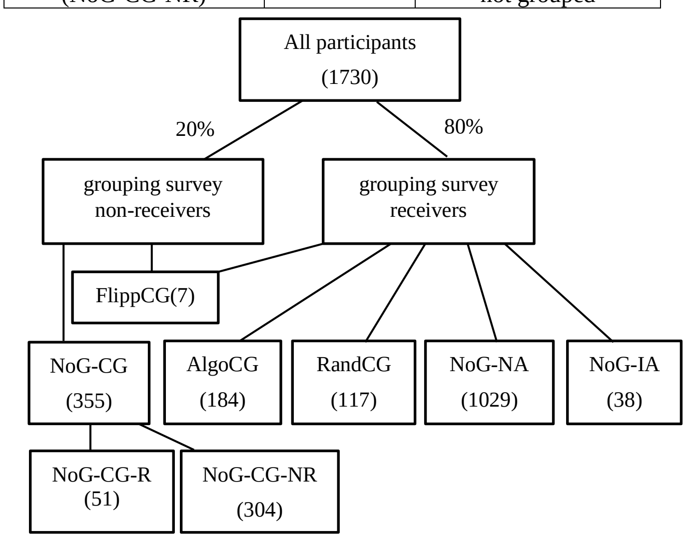

Figure 1. Student conditions with participation numbers.

As a last grouping related intervention, we sent a post grouping survey by the end of the course. This survey was only sent to the 80 percent who had also received the initial grouping survey and contained questions about satisfaction with and intensity of the group work.

3. GROUPING ALGORITHM

In order to create algorithm composed groups (AlgoCG), we used the collected responses from the grouping survey. We first segmented the respondents into five classes according to their collaboration preferences, namely local, email, Facebook, Google+ or Skype. For each class, we extracted each participant's gender, time zone, personality type, learning goal and language for the actual grouping. The task of the algorithm was to compose learning groups consisting of 10 students. Local groups were meant to only contain students from the same cities in order to actually meet up, resulting in very few and small groups qualifying for this option. The main algorithmic challenge was to take into account both heterogeneities (namely gender, personality type and learning goal) and homogeneities (i.e. time zone and language). Concretely, we wanted groups to have e.g. mixed gender, but similar time zone. To solve this optimization problem, we used a k-means clustering algorithm for fixed group sizes, based on [6]. The pseudocode of this algorithm is described in Figure 2 and our implementation in Python is publicly available[1]. In its original form, the algorithm calculates a homogeneity score for a single grouping criterion, like in usual applications of k-means clustering. For our experiment, we modified this algorithm to support multiple criteria and homogeneity as well as heterogeneity at the same time. As a modification, we calculated the group score as the difference between a homogeneity score (on time zone, language and learning goal) and a heterogeneity score (on gender and personality), both of which are actually measured by the Euclidean distance between peers.

```
-----------------------------------------------------------------
Step1: randomly assign students to groups;
Step2: for every group:
            for every student in the current group:
                calculate the possible group scores
                for the student in all the other groups;
                if the student has a higher group score
                in one of the other groups:
                    find the student in the other group with
                    the lowest group score;
                    swap the two students;
Step3: while we are significantly improving the average
       group score, go back to step 2
-------------------------------------------------------------------
```

Figure 2: K-means Clustering for fixed group sizes [6]. Image courtesy: Dirk Uys.

4. EXPERIMENT RESULTS

As a result of our grouping efforts, we composed 22 learning groups in total (4 local meeting groups, 5 Skype groups, 6 Facebook groups, 2 Google+ groups and 5 email groups). The

[1] *https://bitbucket.org/zhilinzheng7/kmeansgrouping*

following sub-sections present three aspects of the experiment results: student engagement, drop-out rate and learning performance.

4.1 Student Engagement

Roughly half of the students enrolled in the course were part of our experiment (1,730 out of 3,209). The other half of the students enrolled after the official start of the course (and, hence, after the start of our experiment), which is a usual pattern for a MOOC.

Overall, the course participants were quite inactive in general, as measured in terms of forum participation. Only 33 students participated in the forum by posting questions, answers or comments. The conducted post grouping survey had nine responses from participants that joined a group. Those respondents spent three hours on average (median one hour) on the group interactions. Further, the Facebook and Google+ groups that were created by us showed some initial greeting messages or comments but no deep, course-related interaction. Hence, our composition did not engage students in collaboration via the social media groups created for that purpose. For other online grouping participants (i.e. email and Skype) who did not answer to our post grouping survey, we cannot make the same conclusion owing to a lack of data.

However, students at least saw small descriptions of their peers in our welcome message and were partly able to see them on social media. Whether this fact, in addition to potentially unobserved interactions (e.g. via email), might have had an impact on the drop-out rate and learning performance, as well as how this relates to survey responsiveness, is analyzed in the following two subsections.

4.2 Drop-out Rate and Survey Responsiveness

We here define a 'drop-out' as any student who did not submit any quiz or assessment, and thereby did not qualify for any course score, after the group assignment.

Figure 3 shows the drop-out rate for all conditions. Unsurprisingly, all seven of the tracked flipped-classroom students stayed in the course (drop-out rate *0%*). In order to test the statistical significance of found differences in the drop-out rates, pairwise z-tests on the different conditions using a two-sided p-value were performed. For our conclusions about significance, we thus applied a Bonferroni correction to the significance level. The p-values in Section 4.2 are given in their non-Bonferroni-corrected form.

First of all, survey responsiveness plays a major role in the analysis. Among the participants of the treatment group that were not grouped, those who gave insufficient survey responses seem to be less likely to drop out than those who did not respond at all, yet this difference is not statistically significant (NoG-IA: *71.05%*, NoG-NA: *82.31%*, *p=0.07*). Further, in the control group without grouping, those who interacted with the motivational survey had a considerably lower drop-out rate than those who did not (NoG-CG-R: *62.75%*, NoG-CG-NR: *82.57%*, *p=0.001*). We can conclude that non-responsive students (with regard to a survey) are more likely to drop out than responsive students.

Hence, when analyzing the interplay between grouping condition and drop-out rate, we need to control for survey responsiveness. Since the randomly composed students did not respond to the grouping survey, we need to compare them to the students in the control group who did not respond to the motivational survey (RandCG: *77.78%*, NoG-CG-NR: *82.57%*, *p=0.26*). And since students from the algorithm composed groups responded to our grouping survey, they need to be compared with the fraction of the control group responding to the motivational survey (AlgoCG: *59.24%*, NoG-CG-R: *62.75%*, *p=0.65*). With this control for survey responsiveness, we thus find no statistically significant effects.

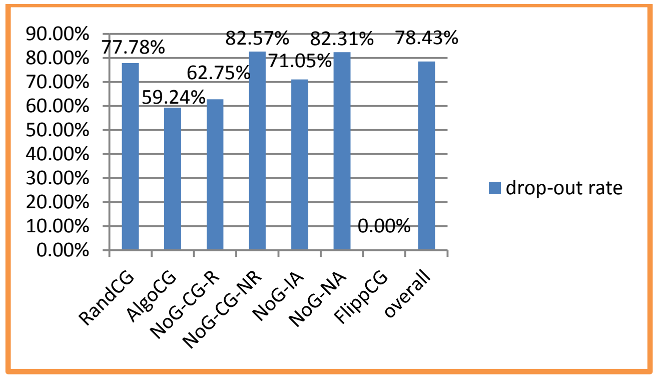

Figure 3: Bar-plot showing student's drop-out rates.

4.3 Learning Performance

In order to analyze the experiment's impact on student's learning performance, we looked at students' scores on quizzes and homework. Figure 4 visualizes average as well as minimum and maximum scores within the various experiment conditions. The flipped classroom condition outperformed all other conditions in terms of median score (FlippCG: 32, others: below 20). However, we do not find evidence for a positive impact of any condition on learning performance as measured by score. A one-way ANOVA implied no statistically significant difference between the conditions (*F(6,518)=1.284, p=0.265*).

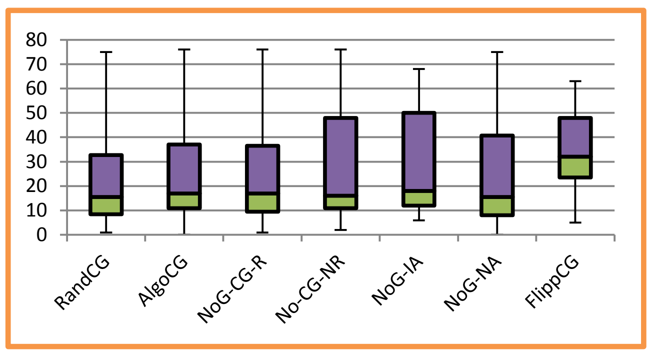

Figure 4. Boxplot showing student's scores from quizzes and homework.

5. DISCUSSION AND FUTURE WORK

In this paper, we presented a scalable and reproducible method to create small groups in online learning environments. We used minimal intervention methods and freely available collaboration tools as well as an adapted k-means clustering algorithm. Within the study, a flipped classroom approach outperformed all composed groups having no drop-out and above average learning scores. This is only partially surprising, as the flipped classroom students were in a formal education setting and most of the others were not. Further, survey responsiveness was found to be predictive of the drop-out rate. Comparing student conditions according to this insight, we found indications that composing small learning groups in MOOCs (at least the way we did it) might not directly increase learning performance online, but could possibly decrease drop-out rates. However, these findings are limited by lack of statistical significance, self-selection biases and little observed interaction in the groups. These limitations need to be addressed within replications and extensions of this experiment.

Statistical significance: The scope of our experiment was a single but massive open online course with quite a high number of participants (1,730), which is far beyond the possibilities of a traditional classroom experiment. However, we faced low response rates and had to assign online students to rather complicated conditions, varying in size between 38 and 1,029 students (cf. Figure 1). The flipped classroom condition only had 7 students. Together, these impediments had a negative impact on the statistical power. For replication, even bigger courses should be chosen.

Self-selection: While only those who completed our grouping survey were assigned to the AlgoCG condition, we chose to compose RandCG from students who did not respond to this survey (for the sake of having enough groups in the AlgoCG condition). This self-selection problem was addressed analytically by also splitting our control group into responders and non-responders to our motivational survey. However, those interventions are not exactly equal: The email containing the motivational survey expresses the wish of the instructor and platform to get to know the students in order to adjust courses accordingly. The email containing the grouping survey, on the other hand, addresses the student's potential wish to collaborate in a group.

Group interaction: Finally, only very low actual collaboration could be observed in the Facebook and Google+ groups. How can small learning groups have an effect if nothing is going on in the groups? Some students claimed in the post grouping survey to have collaborated and it might be the case that the Facebook and Google+ groups were avoided (as an iversity team member was part of the group) and other, private, channels were preferred for collaboration.

In order to overcome the limitations within future student grouping experiments, we deduced new research hypotheses from our results.

Hypothesis 1: Using learning environments that are specifically designed for group work (including reminders, definition of learning goals, assignment of individual group roles or scheduled group meetings) will increase collaboration within small learning groups.

Hypothesis 2: Dynamic group (re-)composition using genetic or particle swarm algorithms will increase collaboration within the small learning groups, by solving the problem of drop-out in learning groups.

Hypothesis 3: Establishing small and regularly interacting sub-communities within a large online course may reduce students' drop-out rate. Just being aware of one another, even if not working together, is crucial.

6. ACKNOWLEDGMENTS

We thank Frank Hoffmann, Michael Sartor and Michael Fröba for collaborating with us and making this research possible; Jo Corrall for helpful feedback; and the China Scholarship Council (CSC) for providing a scholarship that supported this research.

7. REFERENCES

[1] COCEA, M. and MAGOULAS, G., 2010. Group Formation for Collaboration in Exploratory Learning Using Group Technology Techniques. In *Knowledge-Based and Intelligent Information and Engineering Systems*, R. SETCHI, I. JORDANOV, R. HOWLETT and L. JAIN Eds. Springer Berlin Heidelberg, 103-113.

[2] HUANG, Y.M., 2011. A systematic approach for learner group composition utilizing u-learning portfolio. *Journal of educational technology & society 14*, 3, 102-117.

[3] JORDAN, K., 2014. MOOC Completion Rates: The Data. In *Available at:http://www.katyjordan.com/MOOCproject.html*, [Accessed: 27/08/2014].

[4] OH, H., CHUNG, M.-H., and LABIANCA, G., 2004. Group Social Capital and Group Effectiveness: The Role of Informal Socializing Ties. *The Academy of Management Journal 47*, 6, 860-875.

[5] ONAH, D., SINCLAIR, J., and BOYATT, R., 2014. DROPOUT RATES OF MASSIVE OPEN ONLINE COURSES: BEHAVIOURAL PATTERNS. *EDULEARN14 Proceedings*, 5825-5834.

[6] UYS, D., 2014. How we Used the Echo Nest API for Engagement & Learning, Available at: http://info.p2pu.org/2014/01/13/how-we-used-the-echonest-api-for-engagement-learning/, [Accessed: May 2014].

[7] PAPPANO, L., 2012. The year of the mooc. In *In New York Times, November 2012*, Available at: http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html, [Accessed: October 2014].

[8] ROSÉ, C.P., CARLSON, R., YANG, D., WEN, M., RESNICK, L., GOLDMAN, P., and SHERER, J., 2014. Social factors that contribute to attrition in MOOCs. In *Proceedings of the first ACM conference on Learning @ scale conference* (Atlanta, Georgia, USA, 2014), ACM, 2567879, 197-198.

[9] ROSÉ, C.P. and SIEMENS, G., 2014. Shared task on prediction of dropout over time in massively open online courses. In *Proceedings of the 2014 Conference on Empirical Methods In Natural Language Processing Workshop on Modeling Large Scale Social Interaction In Massively Open Online Courses* (Doha, Qata, 2014), 39-41.

[10] SHIMAZOE, J. and ALDRICH, H., 2010. Group Work Can Be Gratifying: Understanding & Overcoming Resistance to Cooperative Learning. *College Teaching 58*, 2, 52-57.

[11] YANG, D., SINHA, T., ADAMSON, D., and ROSE, C.P., 2013. "Turn on, Tune in, Drop out": Anticipating student dropouts in Massive Open Online Courses. In *Proceedings of the Workshop on Data Driven Education, Advances in Neural Information Processing Systems* (2013).