Group Formation for Small-Group Learning: Are Heterogeneous Groups More Productive?

Astrid Wichmann
Educational Psychology Group,
Ruhr-University Bochum
Universitaetsstr. 150
44801 Bochum, Germany
+49 234 3225404
Astrid.wichmann@rub.de

Tobias Hecking
Department of Computer Science and Applied Cognitive Science,
University of Duisburg-Essen
Lotharstraße 63
47057 Duisburg, Germany
+49 203 3791324
hecking@collide.info

Malte Elson
Educational Psychology Group,
Ruhr-University Bochum
Universitaetsstr. 150
44801 Bochum, Germany
+49 234 3226083
malte.elson@rub.de

Nina Christmann
Information and Technology Management,
Ruhr University Bochum
Universitaetsstr. 150
44780 Bochum, Germany
+49 234 3222045
Nina.Christmann@rub.de

Thomas Herrmann
Information and Technology Management,
Ruhr University Bochum
Universitaetsstr. 150
44780 Bochum, Germany
+49 234 3227720
Thomas.herrmann@rub.de

H. Ulrich Hoppe
Department of Computer Science and Applied Cognitive Science,
University of Duisburg-Essen
Lotharstraße 63
47057 Duisburg, Germany
+49 203 3793553
hoppe@collide.info

ABSTRACT
There is an underexploited potential in enhancing massive online learning courses through small-group learning activities. Size and diversity allow for optimizing group composition in small-group tasks. The purpose of this paper was to investigate how groups formed based on learner behavior affect productivity of students in a small-group task. Students classified as high, average and low were randomly assigned to homogeneous or heterogeneous groups. Results indicate that overall, heterogeneous groups were either similarly or a bit more productive than homogeneous groups. Yet, we found that homogeneous groups classified as high-level were as or more than heterogeneous groups. However, heterogeneous groups were still more productive than homogeneous-average and homogeneous-low groups suggesting heterogeneous groups are the best choice for the entire community. Students classified as low-level were more productive in homogeneous groups, suggesting that grouping less active students together, makes social loafing more difficult and students participate more.

CCS Concepts
Applied computing~ Collaborative learning • *Human-centered computing~ Empirical studies in collaborative and social computing*

OpenSym '16, August 17-19, 2016, Berlin, Germany
© 2016 ACM. ISBN 978-1-4503-4451-7/16/08 $15.00
DOI: http://dx.doi.org/10.1145/2957792.2965662

Keywords
Computer-supported collaborative learning, small-group learning, group formation, MOOCs

1. INTRODUCTION
In current practice, large online learning communities such as MOOCs, are mostly geared towards supporting individual, self-paced learning. Instructional videos, literature, as well as quizzes and weekly learning assignments provide the learner with resources and activities, thus fostering learning on the individual level. Although students can usually share information, ask questions, and discuss problems using a threaded discussion forum, only under 5 % of the users take advantage of this [7]. Yet, research on collaborative learning shows that learning in groups can be more effective than working individually [6]. From this point of view, there is an underexploited potential in enhancing massive online learning courses through small-group learning activities. The size and diversity of such communities allows for managing the group composition in such a way as to optimize the conditions for learning.

Collaborative learning can be understood as a situation in which at least two learners learn together [3]. Interacting with others helps making sense of information and building an understanding of concepts to be learned. In MOOCs, collaboration can be an effective means to maintain a high motivation level and overcome difficulties with the learning material [5]. Receiving help from others and discussing problems is important, particularly for students who are motivated to complete the assignments but require assistance. Discussion forums available in MOOCs are not an ideal environment for students to seek help, as they might feel insecure about what question to ask or students wish to remain anonymous [1]. Small-group learning provides a more adequate setting for learners unable to complete assignments on their own. Frequent turn-taking allows for gradually refining the understanding of a problem by just-in-time interactions that are uncommon in asynchronous discussion forums.

However, small-group learning by itself is not always productive. The composition of a group [9] is one factor that affects productivity. MOOCs, consisting of large number of learners, allow for a more productive formation of groups. Participants of MOOCS have culturally diverse backgrounds and enroll in MOOCS for very diverse reasons with diverse sets of skills. The diversity of participants in MOOCs allows forming heterogeneous groups, which might be more productive than homogeneous groups. By taking advantage of available user data, groups can be formed based on individual participation behavior that students exhibit within a MOOC.

2. BACKGROUND

2.1 Small-Group Learning

In this paper, we are concerned with small-group learning, that is, students working together in a small group on a task assignment [3]. In small-group learning, learners have a common goal, which is related to the task provided by the teacher or the facilitator of the learning process. The goal is achieved by shared decision making with shared responsibilities [4]. Learners execute a task collaboratively to be able to complete the task. Thereby students need to engage in coordination processes to make sense of what the goals are and to manage task completion. Coordination with group members requires learners to communicate with each other. In other words, for a successful task execution and coordination learners need to actively participate. Research on collaborative learning has shown that participation rates in online learning settings are rather low and typically characterized by substantial interindividual variability. This is a concern because student participation during collaboration has an influence on student achievement [3]. Low participation of single group members is often described as free rider effect or social loafing: one student participates less and lets others do the bulk of the work [8]. Determining the conditions under which learners engage in social loafing is an important issue. Especially in small-group learning, social loafing can be fatal because it can lead to unsuccessful task completion for all members in the group. In this study, we want to tackle problems in small-group learning. We look at the composition of a group (homogeneous vs. heterogeneous groups) and explore how it affects productivity.

Group composition in collaborative situations constitutes a decisive factor for small-group learning. Heterogeneous group compositions have been discussed to be beneficial for learning in small-group tasks. Researchers used different parameters to form heterogeneous groups.

In this paper, we are interested in how group composition affects productivity of a group. Productivity in small-group learning can be measured in various ways [2]. It can be defined by achievement measured through tests assessing knowledge or different types of skills. Productivity can also be assessed through parameters that are related to the learning process itself. Ye and colleagues [11] looked at students' activities across three MOOCs concerning lecture access and lecture-quiz coverage to better understand and explain drop-out behavior in MOOCs. In this paper, we focus on the learning process, tracking and assessing students' participation patterns during a small-group task to determine productivity.

3. PRESENT STUDY

In MOOCs, small-group tasks provide a good setting for learners to detect and resolve knowledge deficits and cognitive conflicts. Yet, not all small-group learning is productive. In the following study, we investigate the role of group composition in productive small-group learning. Two compositions will be compared: homogeneous and heterogeneous groups. Homogeneity and heterogeneity is established by classifying students as high-level, average-level, and low-level based on student' behavior in the MOOC. Based on group formation research, we assumed that heterogeneous groups are more productive than homogeneous groups. More specifically, we expected that heterogeneous groups exhibit a higher level of student participation than homogeneous groups. Furthermore, we want to explore the role of the individual within a group, and in particular, we are interested if students are as productive if they are assigned to homogeneous groups with other students of the same class compared to heterogeneous groups with students of different classes. Participants

$N = 326$ students from two different universities pursuing various degrees participated in the course for credit. Due to drop-out during the first weeks of the course and people who disapproved the use of their data for research purposes, a total of 120 students remained in our analyses. Permitting student data use was not incentivized with credit points or otherwise.

3.1 Course Description

The course topic was "computer-mediated communication in teaching and learning" and lasted for 14 weeks. It consisted of typical MOOC elements including instructional videos, literature, a discussion forum, and quizzes. In addition, we introduced small-group learning tasks in some of the weeks. The MOOC was offered in the learning environment Moodle[1]. Students were already familiar with Moodle so it made sense to stick with this learning platform. Furthermore, Moodle allowed us to have complete access to user data. For anonymization, we deleted their names and Moodle IDs and replaced them by an extra ID.

The course covered theories of computer-mediated communication and collaborative learning and consisted of 11 thematically distinct course sections. Group formation was applied during one small-group tasks in thematically distinct sections towards the end of the course.

3.2 Small-group tasks

The Small-group task consisted of 29 groups (15 heterogeneous and 14 homogeneous). Students were asked to collaboratively write an essay. They were instructed to draft a text of about 600 words using the writing tool Etherpad (https://en.wikipedia.org/wiki/Etherpad), allowing them to edit the text collaboratively with real-time synchronization. In addition, students used a discussion forum for communication to coordinate the writing process. Students could switch back and forth between Etherpad and the discussion forum.

3.3 Group Formation

In general, assignment to either homogenous or heterogeneous groups was done as follows: Students were first classified in one of the three classes "high", "average", and "low" (Figure 2). Each heterogeneous group included two students classified as "average", one classified as "high" and one classified as "low". Homogeneous groups consisted of either four high-level, four average-level, or four low-level students.

[1] https://moodle.org/

3.4 Parameters to form groups

Basis for forming the groups were parameters capturing individual learner behavior in the MOOC using activity log-files.

For the small-group task, we used forum activity (number of written characters) as a group formation parameter (M=711.61, SD=1648.25, SE=9.93). We classified students as low-level if students either posted nothing or less than 110 characters in the group forum. Students were classified as high-level if students were much engaged in discussions (more than 800 characters).

3.5 Productivity measures

Productivity was measured by students' level of participation during the small group task (see Table 1). Participation was recorded during writing-task using Etherpad (task execution) and activity in group forum (coordination). We considered the following parameters:

Table 1: Participation dimensions.

Writing Tool	
Etherpad: Text-Quantity	Quantity of text measured as number of characters.
Etherpad: Number of Concepts	Total number of domain concepts used in the text.
Discussion Forum	
Forum: Text -Quantity	Quantity of text measured as number of characters. Aggregation per group
Forum: Number of Posts	Number of posts in the group forums.

While most of the participation dimensions described in Table 1 can be calculated straight forward from the log data gathered by the platform, the dimension Number of Concepts required an a priori definition of a domain vocabulary and were calculated semi-automatically. These domain vocabularies specified the most important concepts for the particular course sections in which the group activity took place and were extracted from the provided literature based on manual assessment. Different spellings of the same concept like "cognitive process" and "process of cognition" were mapped to a single concept. In a second automatically executed step, the list of domain concepts was matched by the forum posts and texts in Etherpad in order to calculate the number and coverage of domain concepts present.

4. RESULTS

We will report on productivity within small-group learning looking at the participation dimensions (Table 1). Note that we aggregated participation dimensions Etherpad: Text-Quantity and Forum: Text-Quantity into Etherpad+Forum: Text-Quantity. When planning the small group tasks, we designed the writing tool Etherpad serving as a tool for task execution and the Forum for coordination. Yet, we didn't restrict students to use those tools for respective purposes. As a result, some groups didn't follow those functions we envisioned. A few groups for instance used the forum to draft and edit text using the forum for task execution. Considering this, we looked at text-quantity of Etherpad and Forum in aggregated form.

Small-group task

Taking into account the class level, we found that homogeneous groups classified as high-level were similar productive than heterogeneous groups in Etherpad and the forum, homogeneous-high groups producing a bit more text. Both, homogeneous-high and heterogeneous produced more text than homogeneous average- and low-level groups (Table 2). Surprisingly, homogeneous-high produced only a low number of concepts in Etherpad. Looking at the text quantity overall, homogeneous-high groups seem to be very productive. Checking activity in the forum, we found that several homogeneous-high groups produced text in the forum though not in Etherpad. In other words, they used the forum for task execution instead using the Etherpad as intended, which explains the low number of concepts in Etherpad in the homogeneous-high groups.

Table 2: Productivity during the small-group task.

	Group type	mean	SD
Etherpad+ Forum: Text Quantity	Heterogeneous	9438	4967
	Homogeneous high	10661	6943
	Homogeneous avg.	7623	4748
	Homogeneous low	7692	4044
	Homogeneous all	8651	5660
Etherpad: Number of Concepts	Heterogeneous	14.53	8.8
	Homogeneous high	7.75	6.4
	Homogeneous avg.	12.17	3.76
	Homogeneous low	13.75	7.27
	Homogeneous all	12.49	6.93
Forum: Number of Posts	Heterogeneous	7.5	2.9
	Homogeneous high	10.75	5.9
	Homogeneous avg.	5.3	2.66
	Homogeneous low	4.75	2.75
	Homogeneous all	6.7	4.42

4.1 Participants' productivity based on group assignment

We report effects on productivity depending on the group they were assigned to (see Table 3). For the sake of comprehensibility, we only report results for participants classified as high or low since in these extreme cases, bigger participation changes can be expected than for average students. Productivity was again measured using participation dimensions.

Table 3: Productivity depending on group assignment.

Class	Variable	Homogeneous groups		Heterogeneous groups	
		mean	SD	mean	SD
High	Etherpad+Forum: Text Quantity	2183	2686	3484	2361
	Etherpad: Number of Concepts	7.82	9.35	15.23	13.44
	Forum: Number of Posts	2.53	1.66	2.59	1.7
Low	Etherpad+Forum: Text Quantity	1670	1436	745	1167
	Etherpad: Number of Concepts	4.53	5.03	6.47	11.54
	Forum: Number of Posts	1.20	1.14	0.53	0.53

The results indicate that high-level participants were similarly productive when assigned to heterogeneous in comparison to homogeneous groups across all participation dimensions. During writing, high-level students in heterogeneous groups produced more text in Etherpad and Forum than in homogeneous groups ($F(1,32) = 2.25$, $p = .14$, $eta^2 = .07$). Furthermore, they used more concepts used in Etherpad in heterogeneous groups ($F(1,32) = 3.48$, $p = .07$, $eta^2 = .098$). High-level students made similar number of posts when assigned to heterogeneous or homogeneous groups (Table 3).

By contrast, participants classified as low-level, were more productive in homogeneous groups. During writing, low-level students wrote more text in Etherpad and Forum when assigned to homogeneous groups. Due to the extremely low participation of students classified as low (14 of 15 low-level students in heterogeneous groups had less than 2 forum posts) resulting in extremely skewed distributions of the participation dimensions, non-parametric Kruskal-Wallis test was used ($\chi^2(1)$=4.12, p=0.04). The number of domain concepts used was similar as well as the number of posts both in heterogeneous and homogeneous groups.

5. DISCUSSION

Group constellations are a decisive factor for the productivity within small-group learning. Based on existing group formation research, we assumed that heterogeneously composed groups result in greater productivity. At first glance, we find little evidence for these assumptions in our data. There were overall little differences between heterogeneous and homogeneous groups. Taking into account the class-level, the results convey a more nuanced picture. Homogeneous groups classified as high-level were as much or more productive than heterogeneous groups on most important participation dimensions. Yet, heterogeneous groups were more productive than homogeneous-average and homogeneous-low groups. Furthermore, we found that participants perform quite differently depending on the groups they were assigned to. High-level students seemed to be more productive in heterogeneous groups than in homogeneous groups. They contributed more during the writing task and participated more during discussions. Students classified as low-level, showed quite the opposite behavior. They showed more productivity in homogeneous groups than in heterogeneous groups. In other words, students who were less active (low-level) participated more when being grouped together with other low-level students than being with members who participated more in the forum. One possible explanation is that heterogeneous groups offered more potential for low-level students to engage in social loafing than homogeneous groups. In heterogeneous groups, low-level students can count on productive participants exhibiting high participation to do the work. Our study focused on an important precondition for successful small-group learning that is participation. If we want to instill effective collaboration within MOOC settings, we need to find out which group composition evokes participation best. The results of this study helped us gain a better understanding of productive group compositions and contributed to understanding better how participation in small-groups can be leveraged.

6. REFERENCES

[1] Bull, S., Greer, J., McCalla, G. and Kettel, L. Help-seeking in an asynchronous help forum. In *Proceedings of international conference on artificial intelligence in education*, Seville, Spain, 2001.

[2] Cohen, E. G. Restructuring the classroom: Conditions for productive small groups. *Review of educational research*, 64, 1 (1994), 1-35.

[3] Dillenbourg, P. What do you mean by collaborative learning? *Collaborative-learning: Cognitive and Computational Approaches*, (1999), 1-19.

[4] Dolmans, D. H. and Schmidt, H. G. What do we know about cognitive and motivational effects of small group tutorials in problem-based learning? *Advances in Health Sciences Education*, 11, 4 (2006), 321-336.

[5] Ferschke, O., Howley, I., Tomar, G., Yang, D. and Rosé, C. Fostering Discussion across Communication Media in Massive Open Online Courses. In *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning. (Gothenburg, Sweden)*, 2015, 459-466..

[6] Furberg, A. and Ludvigsen, S. Students' Meaning-making of Socio-scientific Issues in Computer Mediated Settings: Exploring learning through interaction trajectories. *International Journal of Science Education*, 30, 13 (2008), 1775-1799.

[7] Kizilcec, R. F., Schneider, E., Cohen, G. L. and McFarland, D. A. Encouraging Forum Participation in Online Courses with Collectivist, Individualist and Neutral Motivational Framings. In *Proceedings of the European MOOCs Stakeholder Summit*, (Lausanne, Swizerland).

[8] Latane, B., Williams, K. and Harkins, S. Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37, 6 (1979), 822-832.

[9] Webb, N. M., Nemer, K. M. and Zuniga, S. Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Journal, 39,* 4 (2002), 943-989.

[10] Webb, N. M. A process-outcome analysis of learning in group and individual settings. *Educational Psychologist*, (1980) 15, 69–83

[11] Ye, C., Kinnebrew, J. S., Biswas, G., Evans, B. J., Fisher, D. H., Narasimham, G. and Brady, K. A. Behavior Prediction in MOOCs using Higher Granularity Temporal Information. In *Proceedings of the Second ACM Conference on Learning@ Scale.* Vancouver, BC, Canada, ACM, 2015, 335-338.