



# Exploring the impact of chat-based collaborative activities and SRL-focused interventions on students' self-regulation profiles, participation in collaborative activities, retention, and learning in MOOCs

Georgios Psathas<sup>1</sup>  · Stergios Tegos<sup>1</sup> · Stavros N. Demetriadis<sup>1</sup> · Thrasyvoulos Tsiatsos<sup>1</sup>

Received: 15 May 2022 / Accepted: 19 April 2023 / Published online: 23 August 2023  
© The Author(s) 2023

## Abstract

Despite their potential to deliver a high-quality learning experience, massive open online courses (MOOCs) pose several issues, such as high dropout rates, difficulties in collaboration between students, low teaching involvement, and limited teacher–student interaction. Most of these issues can be attributed to the large number, diversity, and variation in self-regulated learning (SRL) skills of participants in MOOCs. Many instructional designers try to overcome these issues by incorporating collaborative activities. Others try to scaffold students' SRL levels by making SRL-focused interventions. However, limited research combines the study of SRL-focused interventions with students' engagement in collaborative activities, course retention, and learning outcomes of MOOC environments. We deployed a programming-oriented MOOC in which we incorporated chat-based collaborative activities, supported by a learning analytics dashboard. Students were asked to complete SRL-focused questionnaires at the beginning and the end of the course. Based on their score, we calculated an average score that forms their SRL level, creating three groups: (a) control, (b) general intervention, and (c) personalized intervention in which we provided personalized interventions. We compared the students' learning outcomes, participation in collaborative activities, and retention in the MOOC. These comparisons provided evidence regarding the positive impact of different intervention modes on students' engagement in collaborative activities and their learning outcomes, with respect to their various SRL profiles. Students allocated to the general and personalized intervention groups displayed increased participation in the collaborative activities and learning outcomes, as compared to students assigned to the control group. We also documented that the SRL interventions positively affected students' course retention.

**Keywords** Massive open online courses · Self-regulated learning · SRL interventions · Collaboration · Collaborative activities

## Introduction

Numerous researchers have described massive open online courses (MOOCs) as one of the most important trends in open online education in the last years (Ebben & Murphy, 2014; Liyanagunawardena et al., 2013; Zhu et al., 2020). MOOCs have been acclaimed for democratizing education and providing individuals access to quality education, regardless of their schedule, geographic location, financial status, spoken language, or academic background (Siemens, 2013). However, although MOOCs offer great learning potential, they also present several issues and challenges that distinguish them from traditional distance learning courses. The most frequently reported MOOC challenges are high dropout rates (Peng & Aggarwal, 2015), difficulties in collaboration between distant students (Smith et al., 2011), and low teaching involvement and teacher–student interaction (Siemens, 2013).

To overcome some of the reported issues, instructional designers have recently begun integrating collaborative activities to enrich MOOCs, thus positively affecting students' engagement, increasing students' interest, reducing dropout, and improving learning outcomes (Kumar & Rosé, 2011; Ortega-Arranz et al., 2017). Szewkis et al. (2011) pointed out six conditions for an activity to be considered collaborative: (a) existence of a common goal (Dillenbourg, 1999), (b) positive interdependence between peers (Johnson & Johnson, 1999), (c) coordination and communication between peers (Gutwin & Greenberg, 2004), (d) individual accountability (Slavin, 1996), (e) awareness of peers' work (Janssen et al., 2007), and (f) joint rewards (Axelrod & Hamilton, 1984).

Students must have the ability to self-regulate their learning to follow the collaborative activities and complete the course. Therefore, they are the ones that decide when and how to interact with the learning material, an autonomy that requires an increased self-regulated learning (SRL) level (Littlejohn et al., 2016; Zimmerman, 2002). In addition, learners must embrace behavioral changes to achieve their goals, which manifests as actions or strategies for persevering until they succeed. Although several researchers have integrated collaborative activities in MOOCs, there is limited research on how teachers can affect students' participation in these activities by making interventions aiming to scaffold students' SRL skills. Targeted SRL interventions can affect students' participation in these activities, retention in a course, and overall learning outcomes in MOOC environments.

Our study tries to fill this research gap by deploying chat-based collaborative learning activities combined with a learning analytics module that provides feedback to peers about their interactions within the collaborative activity of the MOOC. Furthermore, students have the opportunity to reflect on their progress through learning analytics, measure their performance in the chat-based activities, and compare their performance to that of other students. To scaffold students' SRL, we introduced several interventions, and subsequently measured the effectiveness of these interventions on students' participation in collaborative activities, retention, and learning outcomes. Our research aims to evaluate the effect of different interventions on students' self-regulation, participation in collaborative activities, and whether SRL interventions enhance students' learning outcomes and retention in the course.

## Self-regulated learning

SRL includes various cognitive and metacognitive strategies that students need to develop to control and regulate their learning (Pintrich, 1999). According to Zimmerman (1998), self-regulation can be defined as “self-generated thoughts, feelings, and actions for attaining

academic goals.” Students are directed by meta-cognitive reflection on their learning, and they proactively take control of their learning process. For students to succeed in a course, they must have the ability to self-regulate their learning by making decisions regarding how they should interact with the course content and activities.

Self-regulation is even more critical in MOOC settings in order for students to be likely to complete the course. Based on MOOCs open setting and the students’ diversity in SRL skills, it is vital to scaffold students’ self-regulation. This is especially important for students with low SRL skills since many studies have concluded that students who complete the courses present higher self-regulation skills (Reparaz et al., 2020). Terras and Ramsay (2015) provide evidence that students with low SRL scores are prone to drop out of the course. Therefore, it is essential to use tools and methods to support students and scaffold their SRL skills in order to reduce dropout rates and enhance learning performance (Wong et al., 2019).

Since SRL skill improvement can lead to a higher student course completion, many researchers have followed different approaches to positively affect SRL. Some researchers have developed tools to support SRL (Davis et al., 2018; Perez Alvarez et al., 2020). Other researchers have applied different kinds of interventions to scaffold SRL (Jansen et al., 2017). Furthermore, to overcome some of the issues that characterize MOOCs, such as low completion rates and limited teacher–student interaction, instructional designers have recently begun integrating collaborative learning activities to enrich MOOCs (Conole, 2016). Several forms of collaborative activities have been proposed to support students’ learning outcomes and retention. The recent advances in natural language understanding and the positive impact on computer-supported collaborative learning (CSCL) settings have led to an increased application of conversational agents in MOOCs. Such agents aim to simulate the teacher’s behavior by monitoring students’ discussions and delivering prompts to stimulate productive conversational peer interactions, such as argumentation and explicit reasoning (Demetriadis et al., 2018). Thus, the increased results in learning outcomes support the fact that SRL strategies are needed to enhance students’ engagement in collaborative activities. Littlejohn et al. (2016) commented that a positive correlation exists between students’ engagement and learning outcomes.

Järvelä and Hadwin (2013) elaborated on the importance of regulatory processes, and introduced three types of regulated learning that lead to successful collaboration: (a) self-regulation in the group, where individuals in the group regulate their learning; (b) co-regulation, where group members support fellows to regulate their learning; and (c) shared regulation, where the group collectively regulates its learning. To help students reflect on these three dimensions and thus enhance group collaboration, we incorporated a learning analytics (LA) module. Through the visualizations provided by the LA module, students can reflect on their performance at an individual, group level, and course level by comparing their performance to that of other groups.

In our research, we use chat-based activities as the collaborative aspect of the course, complemented by an LA module for individual and group awareness. We apply targeted interventions to foster engagement with the collaborative activities and support students’ SRL skills. Through our study, we want to measure the impact of the SRL-focused interventions on students’ self-regulation profiles, participation in chat-based activities, retention, and learning outcomes in MOOCs.

## Student support and learning analytics

In traditional face-to-face learning environments, teachers reflect on students’ learning and provide individualized support to help them complete the course. In MOOC

environments, such individualized support and guidance may not be easily achieved because of the large number and diversity of students. There is an immense diversity among student profiles in MOOCs. Students with different levels of self-regulation skills, formal education, motivation, and general demographics can enroll in an online course. Thus, students need to be able to self-regulate, take responsibility for their learning, and reflect on their learning process.

According to Zimmerman (2000), learning is a cyclical process divided into three phases: (a) forethought, (b) reactions or performance, and (c) reflection. To provide adequate support to students in MOOCs, instructional designers can use online tools such as LA to support students in this cyclical learning process. Students can potentially use LA to reflect on their performance, compare their performance with peers, and make decisions regarding their learning. According to Lim et al. (2019), individualized support based on LA leads to better learning outcomes. The LA tool can also be used by the instructors to make decisions about which interventions they need to make to scaffold students' learning.

MOOC platforms provide an enormous amount of data from students' interaction with the course environment, the learning components, and the learning activities. To explore the data collected in online learning environments, instructors and learners should have access to a dashboard. LA dashboards allow users to access several visualizations to help them reflect on their learning performance. Self-monitoring provided by dashboard visualizations can be vital for students as they adjust their actions using SRL strategies. Through the LA dashboard, students are able to monitor and reflect on their performance in a collaborative activity at an individual level and compare their performance with their peers in the same or other groups. Based on the information provided by the LA dashboard, they might further decide whether they will change any learning strategy to improve their performance. According to Webber et al. (1993), students engaged in self-monitoring achieve better learning results.

Our study uses chat-based collaborative learning activities and a learning analytics dashboard to provide feedback to peers about their peer-agent interaction in a programming MOOC to support their learning process according to the three phases of Zimmerman's model. Through the learning analytics dashboard, we expect students to reflect on their progress and measure their performance in the chat-based activities during the performance phase. We also expect students to compare their performance with other students who have completed the same activities to help them regulate their learning during the reflection phase. The learning analytics dashboard provides several visualizations that students can use to reflect on their performance.

In addition, we offer several SRL interventions to scaffold students' SRL, enhance their judgment concerning their performance, and measure the impact of SRL interventions on participation during collaborative activities. Through our interventions, we aim to support students in deciding which learning analytics visualization provides them with the appropriate information and how they can interpret this information in alignment with their objective as an aid to help students take control of their learning for better learning outcomes. Aguilar et al. (2021) found positive effect on students' performance when they compared their performance with that of their peers. In addition, Duan et al. (2022) found a positive correlation between students accessing the learning analytics dashboard and their course performance. Based on previous findings, we aim through the interventions to motivate students to compare their performance with their peers' performance by providing insights from the corresponding learning analytics visualizations.

## Rationale and research questions

We deployed the “Programming for non-Programmers” MOOC, which incorporated a series of collaborative chat activities. In addition, to help students self-monitor their learning and reflect on their and peers’ performance in the collaborative activities, we granted them access to an LA dashboard. Through this dashboard, students were able to compare their results with those of others and determine whether they needed to adapt their learning process in order to achieve better learning results.

Several discussion practices have been followed in classroom settings to effectively promote learning and student participation. These practices form the academically productive talk (APT) framework that describes that students should: (a) listen to and build on other students’ contributions, (b) support their contribution to the discussion, and (c) reasonably explain their arguments (Resnick et al., 2010). Considering that academically productive talk is expected to enhance learning outcomes in a collaborative learning setting (Tegos et al., 2016), we aimed to explore whether we could increase students’ participation in collaborative activities via a series of SRL interventions. Drawing on researchers that have previously followed various approaches to scaffold SRL behaviors, this study employed two modes of SRL interventions: (a) a general intervention and (b) a personalized intervention. Both interventions aimed to make students aware of the existence of the LA dashboard and to help them interpret the results both at an individual and a group level. The main difference between these two intervention modes was the personalization of the interventions applied to students in the personalized intervention group. The high-level goal was to investigate whether different SRL intervention modes differentially influence students’ SRL profiles, retention, participation in collaborative activities, and learning outcomes as compared to that of the control group where no interventions were applied.

Under this prism, the research questions that the current study poses are:

- RQ1: To what extent do the different modes of SRL interventions affect students’ SRL profiles?
- RQ2: To what extent do the different modes of SRL interventions affect participation in collaborative activities?
- RQ3: To what extent do SRL profiles predict participation in collaborative activities?
- RQ4: To what extent do the different modes of SRL interventions affect students’ dropout rate?
- RQ5: To what extent do the different modes of SRL interventions affect learning outcomes?

Overall, in view of the SRL theory, the study aims to explore whether SRL interventions affect students’ SRL profiles as well as how a student’s SRL profile correlates with the course retention rate and the learning outcomes.

## Method

### Context

The experiment was conducted in two different “Programming for non-Programmers” MOOC runs. The MOOC introduced students to the basic concepts and structures of algorithmic thinking and familiarized them with the Python programming language. The course consisted of five main modules: (a) computational thinking algorithms and data types, (b) program flow control, (c) data structures, (d) functions and code organization pseudocode, and (e) files, each of them was made accessible to the learners every week.

The modules included video lectures, mini-quizzes, weekly quizzes, weekly assignments, and chat activities, which were incorporated in order to further enhance the learning experience. As shown in Fig. 1, students were required to watch a set of prerecorded video lectures in each module. Following each mini lecture of the “Programming for non-Programmers” course, learners were required to complete a series of multiple-choice questions.



Fig. 1 “Programming for non-Programmers” course structure

Students were informed that they would be able to obtain a course completion certificate if they received at least 60% (six out of ten) in their course grade. Students' course grades depended on their average scores in the mini quizzes and weekly quizzes, and weekly programming assignments, which were multiplied by a fixed coefficient based on the number of chat activities students participated in. The coefficient was 0.8 for students who did not participate in any chat activity, 0.9 for students who participated in one or two activities, and 1 for students who participated in more than three activities. The exact score calculation formula was presented and explained to all students before enrolling in the course and was constantly accessible in the Frequently Asked Questions (FAQ) section.

## Participants

In the "Programming for non-Programmers" MOOC, 3244 students enrolled. Of the total number of students registered, 2409 (74.26%) started working with the course material. We defined a user who started to be a student who had engaged with any of the course material released in the first week. Considering the number of students who started working with the course material and the number of students who completed the course, the completion rate was 52.46%.

To answer our research questions, we considered a subgroup of students to investigate our research questions further. This subgroup consisted of 1807 students who had started working with the course material and completed the intro questionnaire. The rationale for this decision was that we wanted to have participants' SLR profiles to investigate our research questions.

Regarding gender, the distribution was very balanced. More specifically, the student group that reported their gender ( $N=1800$ ) consisted of 867 females (48%) and 933 males (52%). There was a small number of participants ( $N=7$ ) that did not respond to this question.

Most of the students reported their employment status as being unemployed. More specifically, 1100 out of 1794 reported their status as unemployed (61.32%), and 694 (38.68%) indicated their status as currently employed.

## Procedure

After students had registered for the course, they were encouraged to participate in the course-entry questionnaire, which collected data regarding basic demographic information, prior domain knowledge, and SRL score. Learners' SRL score was measured with the Self-regulated Online Learning Questionnaire Revised (SOL-Q-R) translated into Greek, which has been tested for validity in online environments (Jansen et al., 2017). This questionnaire is an improved version of the SOL-Q questionnaire, proven to measure learners' SRL levels in online settings. The SOL-Q-R questionnaire includes 42 questions, divided over seven SRL strategies (details in Sect. 2.5.1). Questions are answered on a seven-point Likert scale, ranging from "not at all true for me" (1) to "very true for me" (7).

We calculated the average SRL score based on students' SRL profile from the course-entry questionnaire, from which we created four clusters. We created four clusters to include a comparable number of students in each cluster, as most students presented mid-level and good self-regulation, and we wanted to have a more equal distribution of students in each cluster. We refer to these clusters as cluster 1, cluster 2, cluster 3, and cluster 4. The distribution of students among the four clusters can be found in Table 1 below.

**Table 1** Analysis of SRL clusters based on SRL score

Cluster	SRL score	SRL profile	<i>N</i>
1	<4	No-to-low self-regulation	178
2	4–5	Mid-level self-regulation	759
3	5–6	Good self-regulation	724
4	>6	High self-regulation	145

The module's material contained the course mini lectures, mini quizzes, the weekly quiz, and the programming assignments made available to the students at the beginning of each week. In addition, we scheduled weekly activities to foster student collaboration.

Considering the conditions of Szewkis et al. (2011) mentioned in the introduction, we embedded chat-based collaborative activities empowered by a conversational agent (Demetriadis et al., 2018), which were expected to promote productive peer dialogue (Tegos et al., 2019). Each chat activity introduced students to an open-ended question (debate), asking them to collaborate in dyads to submit a joint answer as a team (Fig. 2) to fulfill the existence of a common goal and the positive interdependence collaboration condition. As the activity was chat based, students exchanged written messages through the chat activity interface and were encouraged to pay attention to the conversational agent messages ("TIM" in Fig. 2) that may appear during their peer discussion. To complete these activities, students needed to work in dyads and answer the chat topic collaboratively to receive the joint reward that was a grade for the activity. Students were advised to finish studying the respective module material before participating in the collaborative chat activity at the end of each week.

The pairing was based on a first-come-first-served basis. Students visited the activity and waited until another peer connected to the same activity to start working. To help students'

The screenshot displays a chat interface for a collaborative activity. On the left, a sidebar shows the activity title 'TIM' and a 'GET HELP' button. Below this, a 'TOPIC:' section contains a detailed problem statement in Greek about writing code to translate Greek words to English. At the bottom of the sidebar, there is a 'Write your answer to the topic here' text area, a 'Submit' button, and logos for the Erasmus+ Programme of the European Union.

The main chat area shows a conversation between two students, Sophia and Georgia, and the agent TIM. The messages are as follows:

- Georgia: "Does it use ASCII table?" (10:32:48)
- Sophia: "What are the advantages?" (10:33:14)
- Georgia: "To implement the solution with dictionary?" (10:34:16)
- TIM: "You mention the dictionary. Consider this: If you want to locate data in a dictionary with 100 pairs, will the detection time be doubled if your dictionary had 200 pairs? If it had 2 million pairs? What conclusion do you make?" (10:34:17)
- Georgia: "No, it won't double" (10:32:48)
- Sophia: "Tracking time will remain the same" (10:34:53)
- TIM: "Are you talking about the code execution time? Think of this: the user enters alphanumeric data and the code must determine if these data consist a real word in the Greek language. When will this process run faster: when the words are stored in a list or in a dictionary? Please elaborate on your answer." (10:34:53)

At the bottom of the chat area, there is a text input field labeled 'Type your message here' and a 'Send' button.

**Fig. 2** Example from the chat-based activities with agent's intervention (TIM) [Translated from Greek]



pairing, we encouraged them to participate in these activities at specific timeslots every day. Also, we advised students to use the dedicated course forum to post a pairing request. Students could access each chat more than once to complete the activity. In addition, to satisfy individual accountability and awareness of their peers work conditions, we deployed the LA module that was constantly available to students. Through the LA module, students could reflect on their individual and partner’s performance and compare their metrics with all other students that completed the same activity. Based on the information students receive from the LA module graphs, students were able to choose whether to change any learning strategy in order to improve their performance. All student interactions with the LA modules were recorded in the course log files and made available to students through several graphs (Fig. 3).

We allowed all students to participate in the course content following their pace during the first week of their studies. At the end of each week, we downloaded all the students answers to the programming assignment. An automated Python code grader calculated the assignment score with the appropriate feedback. Finally, the scores with the feedback on students’ submissions were uploaded to students’ profiles.

In week 2, we sent a general notification to all participants to inform them about the LA module’s existence, how to access the module, and what information they could get. For the remaining 3 weeks, we provided feedback to students based on their performance during chat activities. In addition, we audited students’ LA data through the corresponding module. Based on the findings, we supported students through appropriate interventions to help them reflect on their performance in the chat activities.

We utilized a final quiz, analyzed in Sect. 2.5.2, to measure the learning impact of the chat-based activities and different SRL interventions on students’ performance and SRL profile. This quiz was made available to all students after they had completed the course.

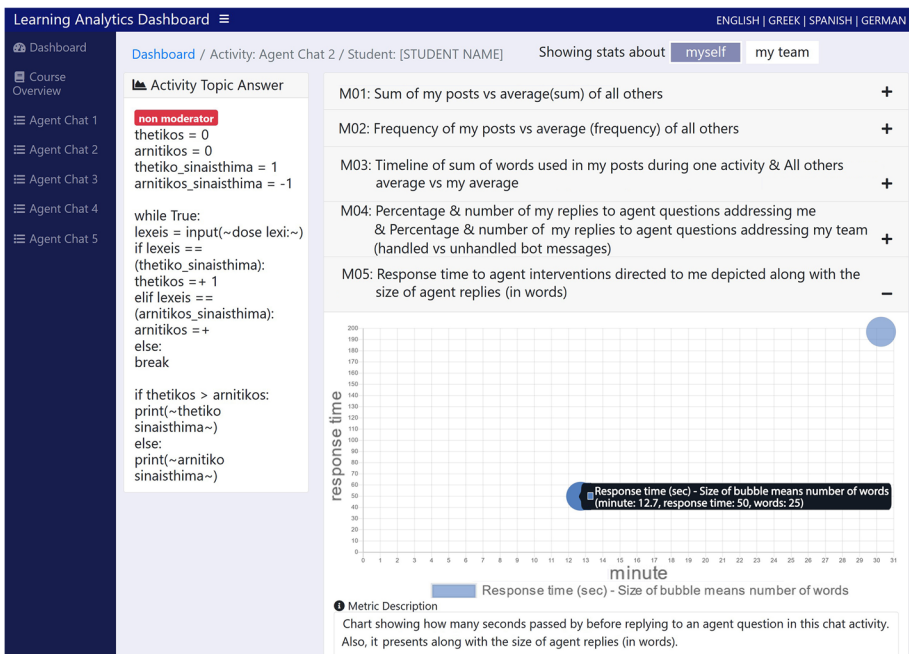


Fig. 3 Sample graph (Graph M05) in the learning analytics module

We advised students to participate in the quiz without referring to the lecture resources (videos, extra material, quizzes) so that it was possible to assess their recall of the material presented in the course. Participation in the final quiz was mandatory for all students wishing to receive a course completion certification.

## Conditions

At the beginning of the third week, we created three groups of students: (a) the control group, (b) the general intervention group, and (c) the personalized intervention group. We randomly selected representatives from all clusters to form the control, the general intervention, and the personalized intervention groups. All groups consisted of 120 students with comparable representatives from each cluster. The distribution of the student's SRL cluster among all groups is depicted in the following Table 2.

### Control group

Out of the 1807 students, we randomly selected 120 to form the control group. The teacher provided basic mobilization to students in this group to access the learning analytics module. This group received only the general email notification intervention at the beginning of week 2. After receiving the general notification, we let the users decide whether to use the learning analytics module to improve their performance or not.

### General intervention group

The general intervention group also consisted of 120 students, including representatives from each cluster, with intense student prompting by the teacher. Students in this group

**Table 2** SRL cluster descriptive statistics and main demographics for representatives in each group

Groups	SRL clusters				<i>N</i>
	1	2	3	4	
<b>Control</b>					
<i>N</i>	25	38	33	24	
Males/females	12/13	15/23	9/24	13/11	120
M	3.38	4.64	5.51	6.36	
SD	0.74	0.25	0.24	0.25	
<b>General intervention</b>					
<i>N</i>	23	40	35	22	
Males/females	13/10	20/20	18/17	10/12	120
M	3.35	4.47	5.39	6.25	
SD	0.64	0.30	0.29	0.21	
<b>Personalized intervention</b>					
<i>N</i>	28	37	33	22	
Males/females	15/13	20/17	16/17	10/12	120
M	3.57	4.53	5.32	6.35	
SD	0.42	0.28	0.21	0.27	

received the general email notification at the beginning of week 2. Starting from week 3 and continuing on a weekly basis until the end of the course, we conducted another three interventions through email messages. All interventions were sent at the beginning of each week. Interventions sent to this group aimed to provide students with information regarding the skills they needed to develop to improve their performance in chat activities and prompted them to elaborate on consideration of which of the graphs presented with the learning analytics offered them valuable information related to these skills. All students in this group received the same exact intervention without any personalization. Each intervention was completely independent of any of the other interventions.

### **Personalized intervention group**

The personalized intervention group consisted of 120 students with representatives from each cluster, similar to the general intervention group, with personalized prompting of students by the teacher. Students in this group received the general email notification at the beginning of week 2. During the whole course, and after the week 2 intervention, we performed another three weekly interventions through email messages sent at the beginning of each week.

The intervention aimed to help students improve their performance in chat-based activities, similar to the interventions made to the general intervention group. There were three main differences between the interventions made in the general intervention group and those made in the personalized intervention group. The first difference was that the interventions applied to the personalized intervention group were more personalized. We included the student's name at the start of the intervention text so that each intervention was addressed to each student independently. The second difference was that each intervention contained information regarding the student's performance based on the previous intervention. In the main body of the intervention, we referred to whether students had complied with our previous recommendations as well as commenting on the results. Finally, the third difference was that we mentioned in the intervention how they could improve specific skills and which graph they should refer to in order to retrieve the relevant information. We connected the graph and the interpretation of the graph, and we kept track of whether they complied or not with our suggestion to comment on our subsequent intervention.

### **Data collection and analysis**

#### **Dependent variable: Student SRL level**

Students were asked to complete the course entry questionnaire containing demographic information (age range, gender, education, employment status), previous MOOC experience, programming languages in general, the Python programming language, and SRL-related questions. During the students first visit to the course, they were informed of the data collection procedure and were asked to consent before starting to work with the course material. In the course entry questionnaire, we incorporated the SOL-Q-R questionnaire to record students' respective SRL profiles. Table 3 indicates the seven SRL strategies that the SOL-Q-R includes, the number of questions in each strategy, and an indicative item.

Subsequently, after finishing all the course activities, the students completed another questionnaire, namely the course exit questionnaire. The course exit questionnaire contained an evaluation of the quality of the course, the instructors, and the chat activities. The same SOL-Q-R questionnaire was also included in this survey to calculate the SRL difference and analyze the impact of the SRL intervention on students.

**Table 3** Indicative item for each SRL strategy from SOL-Q-R questionnaire

SRL strategy	Number of questions	Indicative item
Metacognitive activities before learning containing	7	I set specific goals before I begin a task in this online course
Metacognitive activities during learning	7	When I study for this online course I try to use strategies that have worked in the past
Metacognitive activities after learning	6	I think about what I have learned after I finish working on this online course
Time management	5	I allocate studying time for this online course
Environmental structuring	4	I choose the location where I study for this online course to avoid too much distraction
Persistence	7	Even when I feel lazy or bored when I study for this online course, I finish what I planned to do
Help seeking	6	When I do not fully understand something, I ask other course members in this online course for ideas

## Dependent variable: Learning outcomes

To measure students' learning outcomes, we constructed a course exit quiz. The course exit quiz consisted of 20 multiple-choice questions: 10 related to the topics discussed during the chat activities, and 10 related to concepts discussed during the course video lectures. There was no time limit on answering the quiz, and each student was only able to participate in the quiz once. We collected all answers to the quiz, and categorized them according to whether they were correct or not.

## Dependent variable: Student participation

As shown in Fig. 1, for students to complete the course, they had to complete several weekly activities, such as mini quizzes, weekly quizzes, and programming assignments. All students' access and scores in these activities were stored on the platform's course log files. Also, course log files stored information regarding whether students tried to participate in the chat activities, the status of these activities (completed or unable to find a partner), and the total number of visits to the chat activities learning analytics module. We used these metrics to evaluate access and scores in the course activities in order to compute the students' level of engagement in the course activities as well as the dropout rate.

## Data analysis

In the current study, we leveraged data from the course entry questionnaire, course exit questionnaire, course exit quiz, and course logs to answer the research questions. To measure students' SRL score, we first calculated the Cronbach's Alpha reliability test to validate the answers to the SRL profile related questions from the course entry and course exit questionnaires, in order to check internal consistency. The Cronbach's Alpha was calculated for the each entire SRL questionnaire, and the mean and standard deviation measures were calculated for each student for each of the seven SRL strategies independently. Based on the average SRL score from the course entry questionnaire, we classified students into one of the four SRL clusters. We randomly picked representatives from each cluster to form the condition groups, trying to keep the same number of representatives from each cluster.

We compared students' SRL profiles reported in the course-entry and course-exit questionnaires respectively. Then, we applied analysis of covariance (ANCOVA) to study whether there were differences in students' SRL both at the start and the end of the course, to answer RQ1. Following the clustering procedure, we retrieved the number of times students participated in the following chat-based activities after each intervention from the course log files to elaborate on RQ2.

We combined students' SRL data from the course intro questionnaire with the participation information to evaluate how different SRL interventions affected student participation in the collaborative activities in order to answer RQ3. We used linear regression to understand the effect of the different interventions made to each group on student participation in the collaborative chat activities.

We also researched whether there is a link between the students' condition group and student dropout rates in order to elaborate on RQ4. To answer that question, we conducted a chi-square test of homogeneity between students' condition groups extracted from the SRL-related answers on the entry questionnaire, and the course logs. We wanted to investigate whether the intervention in general, or the kind of intervention in particular, are significant factors determining student learning outcomes. In accordance with RQ5, we sought

to elaborate on the impact on students after receiving different interventions, in terms of the course final quiz. We conducted a one-way ANOVA incorporating students' condition group and the scores they received in the final quiz to answer this question.

## Results

### RQ1: To what extent do the different modes of SRL interventions affect student SRL profiles

We calculated each student's SRL level using all SRL-related data from the course entry and course exit questionnaires. Analyzing the data we collected from the two questionnaires, we explored the SRL profiles of students attending their respective programming MOOC. Table 4 below presents descriptive statistics for the SRL score

**Table 4** SRL strategy descriptive statistics

SRL strategy score per group	Course entry questionnaire ( $N=360$ , $\alpha=0.958$ )				Course exit questionnaire ( $N=277$ , $\alpha=0.927$ )			
	M	Min	Max	SD	M	Min	Max	SD
Metacognitive activities before learning								
Control group	5.21	1	7	1.23	4.98	1	7	1.32
General intervention group	5.12	1	7	1.36	5.05	2	7	1.08
Personalized intervention group	5.06	2.57	7	1.18	4.88	1.71	7	1.19
Metacognitive activities during learning								
Control group	4.99	1	7	1.20	4.93	1	7	1.40
General intervention group	4.86	1	7	1.19	4.78	2.29	7	0.97
Personalized intervention group	4.87	1	7	1.23	4.82	1	7	1.04
Metacognitive activities after learning								
Control group	5.22	1	7	1.38	5.05	1	7	1.53
General intervention group	5.04	1	7	1.38	4.88	1.67	7	1.25
Personalized intervention group	5.05	1.50	7	1.26	4.83	1	7	1.23
Time management								
Control group	4.65	1	7	1.05	4.44	1	7	0.97
General intervention group	4.58	1	7	1.02	4.47	2.40	6.40	0.70
Personalized intervention group	4.60	1	7	0.99	4.55	2	7	0.85
Environmental structuring								
Control group	5.52	1	7	1.52	5.40	1	7	1.57
General intervention group	5.57	1	7	1.40	5.59	1.50	7	1.47
Personalized intervention group	5.33	1	7	1.51	5.53	2	7	1.37
Persistence								
Control group	4.68	1	7	1.46	4.55	1	7	1.46
General intervention group	4.40	1	6.86	1.40	4.43	1.57	7	1.41
Personalized intervention group	4.46	1	7	1.35	4.70	1.71	7	1.25
Help seeking								
Control group	4.52	1	7	1.60	2.69	1	7	1.69
General intervention group	4.55	1	7	1.75	2.90	1	7	1.61
Personalized intervention group	4.74	1	7	1.48	3.67	1	7	1.61

for each condition and for each SRL strategy. The entry and exit questionnaires exceed the 0.70 value for alpha, which is the minimum acceptable value, indicating the strong internal validity of the questionnaires. The course entry questionnaire provided an overall Cronbach's alpha  $\alpha=0.958$ , and the course exit questionnaire had an overall Cronbach's alpha  $\alpha=0.925$ .

Subsequently, we recorded the SRL level difference for each student between self-reported SRL-related questions from the course entry and course exit questionnaires to analyze the impact of the SRL interventions on students. An ANCOVA was run to determine the effect of the three study conditions on post-course SRL scores after controlling for the students' SRL levels reported before the beginning of the course. After adjustment for pre-course SRL scores, there was a statistically significant difference in students' SRL levels between the conditions,  $F(2, 273)=5.082$ ,  $p<0.01$ , partial  $\eta^2=0.036$ . Adjusted mean is presented unless otherwise stated. The students' SRL level was statistically significantly greater in the personalized intervention group ( $M=4.683$ ,  $SE=0.068$ ) as compared to the control group ( $M=4.355$ ,  $SE=0.077$ ), with a mean difference of 0.328, 95% CI 0.079, 0.577,  $p<0.01$ . There was no statistically significant difference in students' SRL levels between the control and the general intervention groups or between the general and personalized intervention groups.

### **RQ2: To what extent do the different modes of SRL interventions affect participation in collaborative activities?**

We audited the relevant course log files and extracted all data related to student participation in the collaborative activities. A Kruskal–Wallis  $H$  test was run to determine the differences in student participation in collaborative activities between the groups of students in the three study conditions. Distributions of student participation scores were similar for all groups, as assessed by visual inspection of a boxplot. Median CWWS scores were statistically significantly different between groups,  $H(2)=83.541$ ,  $p<0.01$ .

Pairwise comparisons were conducted using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted  $p$ -values are presented. This post hoc analysis indicated statistically significant differences in student participation between the control (mean rank = 116.70) and the general intervention condition (mean rank = 204.18;  $p<0.01$ ), as well as the control and the personalized intervention condition (mean rank = 220.62;  $p<0.01$ ), but not between other group combinations.

### **RQ3: To what extent do SRL profiles predict participation in collaborative activities?**

With respect to RQ3, a linear regression was run to understand the effect of students' SRL level on their participation in collaborative activities. To assess linearity, a scatterplot of students' participation against SRL level with a superimposed regression line was plotted. Visual inspection of these two plots indicated a linear relationship between the variables. The prediction equation was as follows: students' participation in collaborative activities =  $2.794 + 0.201 \times (\text{SRL cluster})$ . The regression analysis revealed that the students' SRL level could statistically significantly predict students' participation in collaborative activities,  $F(1, 358)=6539$ ,  $p<0.012$  and SRL level with Cohen's  $f=0.12$ , a small size effect according to Cohen (1988).

#### RQ4: To what extent do the different modes of SRL interventions affect student dropout rate?

To elaborate on RQ4, we extracted information from the course activities log files to measure student engagement with the course activities, and we correlated these measurements with the study conditions. Out of the 360 students that formed the three groups, it was found that 84 (23.3%) of the students dropped out and did not complete the course. While further investigating whether there is an association between the study conditions and the student dropout rates, a chi-square test of homogeneity revealed a statistically significant difference between the three conditions ( $\chi^2 = 22.08$ ,  $p < 0.01$ ). The results for the students in all conditions are tabulated in Table 5 below. More specifically, following the personalized intervention, 87.5% of the students continued in the course compared to 80.0% of the students in the general intervention condition and 62.5% in the control condition.

Post hoc analysis involved pairwise comparisons using the  $z$ -test of two proportions with a Bonferroni correction. The proportion of students classified as “retained users” after the personalized intervention was statistically significantly higher than that reported in the control condition,  $p < 0.01$ . Likewise, the proportion of students classified as “retained users” following the general intervention was also significantly higher than in the control condition,  $p < 0.01$ . The proportion of the students that dropped out in the two treatment conditions was not found to have a statistically significant difference.

#### RQ5: To what extent do the different modes of SRL interventions affect learning outcomes?

To answer RQ5, we used the data from the course exit quiz. A one-way ANOVA was conducted to determine whether the score in the course exit quiz was different for the three groups that received different SRL interventions, i.e., the control group, the general intervention group, and the personalized intervention group.

Data are presented as mean  $\pm$  standard deviation. The score in the course exit quiz was statistically significantly different between the groups that received different interventions.,  $F(2, 357) = 13.650$ ,  $p < 0.01$ , partial  $\eta^2 = 0.071$ . Scores in the final quiz were higher for students from the general intervention ( $M = 11.46$ ,  $SD = 6.864$ ) and personalized intervention ( $M = 11.71$ ,  $SD = 6.129$ ) groups as compared to the scores of the students from the control group ( $M = 8.46$ ,  $SD = 7.666$ ). Tukey post hoc analysis revealed that the mean increase in the scores for students in the general intervention

**Table 5** Student drop out rate per condition

		Conditions groups			Total
		Control	General intervention	Personalized intervention	
Course participation	Retained	75 (62.5%)	96 (80.0%)	105 (87.5%)	276 (76.7%)
	Drop out	45 (37.5%)	24 (20.0%)	15 (12.5%)	84 (23.3%)
	Totals	120 (100%)	120 (100%)	120 (100%)	360 (100%)



group (3.74, 95% CI 1.80, 5.68) as compared to the scores of the students from the control group was statistically significant ( $p < 0.01$ ). Also significant was the increase of the scores for students in the personalized intervention group (3.72, 95% CI 1.78, 5.66,  $p < 0.01$ ) compared with students from the control group, but no other group differences were statistically significant.

## Discussion

Given the peculiarities that characterize MOOCs, such as the openness and easy access to everyone, students are required to develop specific skills in order to be likely to complete a course. Self-regulation is regarded as one of the most critical aspects of a students' learning process. Based on the SRL existence necessity, this work presents a study investigating the impact of different SRL interventions on students' behavior in a MOOC, which were designed to scaffold students' SRL strategies. The study results are quite promising and indicate the importance of delivering interventions that target student self-regulation.

Self-reported SRL scores had been distributed across the whole SRL-score scale, showing heterogeneity, with many students in either category representing high and low scores. Therefore, we delivered a series of SRL interventions to support students' self-regulation profiles, drawing on the general and personalized SRL intervention modes used in this study. Given that the only difference between the general and the personalized intervention groups was personalization, it was investigated whether these different SRL interventions would be able to significantly impact students' self-regulation profiles.

To answer RQ1, we measured the differences in students' SRL profiles by comparing self-reported SRL data from the course-entry and course-exit questionnaires. Our findings align with previous studies (Jansen et al., 2017; Lim et al., 2023), which pointed out significant differences in SRL skills for the students who complied with the interventions. In addition, our analysis showed that even though all groups' average SRL score decreased, students allocated to the personalized intervention group provided a smaller decrease in SRL profile, as measured in the course exit questionnaire. Although students assigned to the control group provided lower levels of SRL, we did not observe a statistically significant difference in the SRL profiles for students allocated to the general intervention group. The general decrease in the SRL profiles might possibly be attributed to most students participating in a MOOC for the first time, so they did not have the experience to properly self-report the SRL questionnaire items. As they kept learning, students engaged in SRL strategies, formed a better understanding, and they evaluated the same SRL items in the course exit questionnaire differently. More specifically, if we compare the SRL levels between the entry and exit questionnaires for each SRL strategy, students assigned to the personalized intervention group provided elevated skills in "Environmental Structuring" and "Persistent" SRL strategies compared with students allocated to the other groups. Moreover, even though the "Help Seeking" SRL strategy decreased in all groups, students from the personalized intervention group provided a smaller decrease, indicating that they were better able to regulate in order to complete the course.

Since the only difference in the intervention received by the personalized intervention group, which provided better SRL scores in the exit questionnaire, was personalization, we concluded that personalization in the SRL interventions plays a significant role in students' SRL scores. That is surprising as, in a 5 week MOOC, students were not expected to make substantial changes in their SRL strategies. Therefore, personalized interventions seem to

be promising when the impact on students' SRL is concerned. Given the open nature of the MOOCs, which tend to attract a large number of students, and the need to scaffold students' SRL skills to help them complete the course, systems that provide personalized interventions to students might prove to be helpful.

With respect to RQ2, it was examined whether the different modes of SRL intervention employed, i.e., the study conditions, might have a substantial impact on student participation in collaborative activities. Davis et al. (2018) provided evidence that interventions enhance engagement with course activities. Our study extends previous research, investigating the effect on participation in collaborative activities. As collaborative activities enhance student learning outcomes, participating in such activities is very important. We tracked student participation in the chat activities through the platform's log files. We observed that students from both the general intervention and personalized intervention groups provided higher levels of engagement in the chat activities than the students from the control group. Most students of the control group participated only in one chat activity, and even the ones that participated in the initial chat activities seemed to quickly lose their interest as the course progressed. In fact, student participation in chat activities for students from the control group was constantly decreasing as the course progressed. On the contrary, the chat participation of the students from the other two groups remained relatively unaffected throughout the course activities. The observed number of students from the general intervention group that did not participate in any chat activity was lower than the number of students from the control group. Even lower numbers were observed in students from the personalized intervention group. It should be noted that the only difference between the control group and the general and personalized intervention groups was that we constantly intervened in their learning process by sending SRL-related information that motivated students to participate more in these activities. More specifically, the interventions demonstrated the benefits of participating in collaborative activities by providing real-world examples, which appeared to slightly boost student motivation. We found that interventions highly impacted student engagement with the collaborative activities. As these differences between general and personalized intervention groups are insubstantial, we concluded that any intervention is sufficient and could lead to an increased participation in chat activities.

Furthermore, in an attempt to address RQ3, we investigated whether student SRL profiles predict their participation in collaborative activities. Previous studies have investigated the contribution of SRL profiles to prediction models related to course success, meaning completion and course grades (Kizilcec et al., 2017; Maldonado-Mahauad et al., 2018; Moreno-Marcos et al., 2020). Interestingly, it was found that students' pre-course SRL levels effectively predict the degree of student participation in synchronous collaborative activities. We consider this important for the purpose of enabling instructors and researchers to employ early detection strategies relating to SRL and in order to develop a prediction model for student participation in the collaborative activities of a course. Although SRL-related information predicts participation, further investigation is needed to define which specific SRL strategy mostly affects this prediction. Nevertheless, this could open up new opportunities for the orchestration of the collaborative learning activities and the design of adaptive collaborative learning support that adjusts automatically based on students' SRL levels.

In an attempt to address RQ4 and RQ5, the SRL intervention modes were also found to have a statistically significant effect on both students' dropout rates and the reported learning outcomes. After comparing student retention rates throughout the MOOC, it was revealed that the portion of students who dropped out of the course was significantly

reduced in the student groups that had received either the general or the personalized SRL interventions. Better retention rates were observed for students from the personalized intervention group, followed by the general intervention group. A significant difference was not found while comparing the two SRL intervention modes, though. This leads us to believe that even the generalized SRL intervention mode, which requires less setup effort on behalf of the teaching staff, can be effective in improving learning outcomes and reducing dropout rates. Based on the literature, the dropout rate in MOOCs can be as high as 93% (Chen & Zhang, 2017; Peng & Aggarwal, 2015), which is a significant concern and the most critical and highest mentioned issue in the literature. Finding ways to reduce this number can play a vital role in the MOOC ecosystem. Regarding the learning outcomes, we observed statistically significant differences between the two intervention groups and the control group. Students from the two intervention groups provided an elevated number of correct answers to the exit quiz. We argued that this happened due to the increased participation in the collaborative activities that led to better learning outcomes.

## Limitations and future research

While we presented some valuable insights regarding student participation in collaborative activities in MOOCs that reported better retention in the course and elevated learning outcomes, there are some limitations in our study. Students that attended the “Programming for non-Programmers” MOOC were located in the same country and speaking the same language, which affects the external validity of our results. Ogan et al. (2015) instruct us that socio-cultural context plays an essential role in how students engage with online courses. Students from different areas with different socio-cultural backgrounds need to be included in our research to better generalize our findings.

Another limitation was the fact that we did not study each intervention strategy’s effect separately. Instead, we calculated the SRL profile as a one-score metric with no qualitative indications of the type of metacognitive skills students mastered. Not including qualitative data might reduce the effect of our findings, especially related to RQ1. The duration of the course can also be considered a limitation. The course we deployed for our research lasted 5 weeks, a relatively small period of time for students to change their SRL strategies.

An additional limitation was that students’ SRL profile were calculated based on self-reported data. Students answered the course-entry and course-exit questionnaires that contained SRL-related questions, and based on their answers, we calculated the average SRL scores. Students participated in the same questionnaire at the beginning and end of the course, which might produce bias. There was no instrument or clickstream data to validate the self-reported SRL measures with their actual activity in the course modules. Also, there was no correlation between students’ prior knowledge and course learning outcomes. In the course exit quiz, we measured student learning outcomes. Still, we did not consider whether students had good knowledge of the course material or whether the better learning outcomes were the effect of the interventions. Further research could also improve the correlation between access to the learning analytics module and participation in the course activities. Even though we manually correlated student performance by elaborating on the relevant graphs to provide them with the corresponding feedback, we could also collect data regarding access to the learning analytics module before and after each intervention.

Finally, in our research, we treated the SRL profile score as an average score that emerged from the course entry questionnaire. A future step would be to analyze which specific SRL strategy of the seven included in the SOL-Q-R questionnaire mostly affects our findings. Knowing which strategies mostly affect the results could make the interventions even more targeted to the specific characteristics. Also, the insights could be more accurate if the self-reported SRL measures were combined with clickstream data or social interactions. For example, self-reported help-seeking SRL levels can be combined with the number of access to the forum or the tools that aim to provide help to students. Analyzing this information, can allow us calculate more accurate differences to SRL-related measures.

## Conclusions and implications

The current study aimed to examine the effects of student's self-regulation profiles on participation in collaborative activities, retention, and learning in MOOCs. Also, through our research, we tried to investigate whether students' SRL profiles could become a predictor of participation in collaborative activities. To this extent, we analyzed data from questionnaires provided to students and log files extracted from the MOOC platform. Based on our findings, participation in collaborative activities, student learning outcomes, and dropout rates can be positively affected by combining collaborative activities and interventions to scaffold students' SRL profile development. Another interesting finding from our study was that students' self-reported SRL profiles predict participation in collaborative activities. Knowing this and understanding how to intervene to support students with lower SRL scores seems to be a critical issue worth further exploration.

More specifically, we found that SRL-related interventions increased student participation in collaborative activities. As collaborative activities have been proved to contribute to better learning outcomes and positively affect retention in a course (Kumar & Rosé, 2011; Ortega-Arranz et al., 2017), it is crucial (a) to implement such activities in MOOCs and (b) to encourage students through interventions to interact with these activities. This combination can potentially elevate both the learning outcomes and the retention rates of a MOOC.

Furthermore, the differences in completion rates between the students from the general and the personalized intervention groups interestingly and clearly showed that any continual feedback results in greater completion rates. Since the most significant issue of MOOCs is the high dropout rate, it is substantial to structure the courses in a way that can increase completion by incorporating collaborative activities in courses.

Several types of collaborative activities can be embedded in a MOOC. For example, our research employed chat-based activities that students had to complete in dyads. We believe that this research opens new opportunities to use SRL-related interventions for both instructors and researchers. If our findings can be generalized to several types of collaborative activities, they may impact the structure of many online courses aiming to reduce dropout rates and enhance learning outcomes.

**Funding** Open access funding provided by HEAL-Link Greece.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aguilar, S. J., Karabenick, S. A., Teasley, S. D., & Baek, C. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. *Computers & Education*, *162*, 104085. <https://doi.org/10.1016/j.compedu.2020.104085>
- Axelrod, R., & Hamilton, W. (1984). The evolution of cooperation. *Science*, *211*(4489), 1390–1396. <https://doi.org/10.1126/science.7466396>
- Chen, Y., & Zhang, M. (2017). MOOC student dropout. *Elektrotechnik Und Informationstechnik*. <https://doi.org/10.1145/3063955.3063959>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Conole, G. (2016). MOOCs as disruptive technologies: strategies for enhancing the learner experience and quality of MOOCs. *Revista de Educación a Distancia (RED)*, *50*, 2. <https://doi.org/10.6018/red/50/2>
- Davis, D., Triglianios, V., Hauff, C., & Houben, G.-J. (2018). SRLx: A personalized learner interface for MOOCs. *Lifelong Technology-Enhanced Learning*, 122–135. [https://doi.org/10.1007/978-3-319-98572-5\\_10](https://doi.org/10.1007/978-3-319-98572-5_10)
- Demetriadis, S., Tegos, S., Psathas, G., Tsiatsos, T., Weinberger, A., Caballé, S., Dimitriadis, Y., Sanchez, E., Papadopoulos, P. M., & Karakostas, A. (2018). *Conversational Agents as Group-Teacher Interaction Mediators in MOOCs*. 2018 Learning With MOOCs (LWMOOCs). <https://doi.org/10.1109/lwmoocs.2018.8534686>
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-Learning: Cognitive and Computational Approaches* (pp. 1–19). Elsevier.
- Duan, X., Wang, C., & Rouamba, G. (2022). *Designing a Learning Analytics Dashboard to Provide Students with Actionable Feedback and Evaluating Its Impacts*. Proceedings of the 14th International Conference on Computer Supported Education. <https://doi.org/10.5220/001116400003182>
- Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252. <https://doi.org/10.1080/00401706.1964.10490181>
- Ebben, M., & Murphy, J. (2014). Unpacking MOOC scholarly discourse: A review of nascent MOOC scholarship. *Learning, Media and Technology*, *39*(3), 328–345. <https://doi.org/10.1080/17439884.2013.878352>
- Gutwin, C., & Greenberg, S. (2004). The importance of awareness for team cognition in distributed collaboration. *Team Cognition: Understanding the Factors That Drive Process and Performance*, 177–201. <https://doi.org/10.1037/10690-009>
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*, *49*(4), 1037–1065. <https://doi.org/10.1016/j.compedu.2006.01.004>
- Jansen, R. S., Van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education*, *29*(1), 6–27. <https://doi.org/10.1007/s12528-016-9125-x>
- Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. <https://doi.org/10.1080/00461520.2012.748006>
- Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory into Practice*, *38*(2), 67–73. <https://doi.org/10.1080/00405849909543834>
- Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. H. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18–33. <https://doi.org/10.1016/j.compedu.2016.10.001>

- Kumar, R., & Rosé, C. P. (2011). Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies*, 4(1), 21–34. <https://doi.org/10.1109/tlt.2010.41>
- Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2019). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*, 101202. <https://doi.org/10.1016/j.learninstruc.2019.04.003>
- Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Rakovic, M., Molenaar, I., Moore, J., & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, 139, 107547. <https://doi.org/10.1016/j.chb.2022.107547>
- Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education*, 29, 40–48. <https://doi.org/10.1016/j.iheduc.2015.12.003>
- Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008–2012. *The International Review of Research in Open and Distributed Learning*, 14(3), 202. <https://doi.org/10.19173/irrodl.v14i3.1455>
- Maldonado-Mahauad, J., Pérez-Sanagustín, M., Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., & Delgado-Kloos, C. (2018). Predicting learners' success in a self-paced MOOC through sequence patterns of self-regulated learning. *Lifelong Technology-Enhanced Learning*, 355–369. [https://doi.org/10.1007/978-3-319-98572-5\\_27](https://doi.org/10.1007/978-3-319-98572-5_27)
- Moreno-Marcos, P. M., Muñoz-Merino, P. J., Maldonado-Mahauad, J., Pérez-Sanagustín, M., Alario-Hoyos, C., & Delgado-Kloos, C. (2020). Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs. *Computers & Education*, 145, 103728. <https://doi.org/10.1016/j.compedu.2019.103728>
- Ogan, A., Walker, E., Baker, R. S., Rodrigo, M. M. T., Soriano, J. M., & Castro, M. J. (2015). Towards understanding how to assess help-seeking behavior across cultures. *International Journal of Artificial Intelligence in Education*, 25(2), 229–248. <https://doi.org/10.1007/s40593-014-0034-8>
- Ortega-Arranz, A., Muñoz-Cristóbal, J. A., Martínez-Monés, A., Bote-Lorenzo, M. L., & Asensio-Pérez, J. I. (2017). Gamifying collaborative activities in MOOCs. In *Proceedings of the Fifth European MOOCs Stakeholders Summit, eMOOCs 2017* (pp. 28–33). Madrid, Spain.
- Peng, D., & Aggarwal, G. (2015). Modeling MOOC Dropouts. *Entropy*, 10(114), 49944.
- Perez Alvarez, R. A., Maldonado, J., Sharma, K., Sapunar, D., & Perez-Sanagustin, M. (2020). Characterizing learners' engagement in MOOCs: An observational case study using the NoteMyProgress tool for supporting self-regulation. *IEEE Transactions on Learning Technologies*, 1–1. <https://doi.org/10.1109/tlt.2020.3003220>
- Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459–470. [https://doi.org/10.1016/s0883-0355\(99\)00015-4](https://doi.org/10.1016/s0883-0355(99)00015-4)
- Reparaz, C., Aznárez-Sanado, M., & Mendoza, G. (2020). Self-regulation of learning and MOOC retention. *Computers in Human Behavior*, 111, 106423. <https://doi.org/10.1016/j.chb.2020.106423>
- Resnick, L. B., Michaels, S., & O'Connor, M. C. (2010). How (well-structured) talk builds the mind. In D. D. Preiss & R. J. Sternberg (Eds.), *Innovations in educational psychology: Perspectives on learning, teaching, and human development* (pp. 163–194). Springer Publishing Company.
- Siemens, G. (2013). Massive open online courses: Innovation in education? In R. McGreal, R. Kinuthia, W. Marshall, S., & McNamara, T. (Eds.), *Open Educational Resources: Innovation, Research and Practice* (pp. 5–15). Athabasca University Press. <http://oasis.col.org/handle/11599/486>
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43–69. <https://doi.org/10.1006/ceps.1996.0004>
- Smith, G. E., Sorensen, C. M., Gump, A., Heindel, A. J., Caris, M., & Martinez, C. J. (2011). Overcoming student resistance to group work: Online versus face-to-face. *Internet and Higher Education*, 14(2), 121–128. <https://doi.org/10.1016/j.iheduc.2010.09.005>
- Szewkis, E., Nussbaum, M., Rosen, T., Abalos, J. L. N., Denardin, F. K., Caballero, D., Tagle, A., & Alcolado, C. (2011). Collaboration within large groups in the classroom. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 561–575. <https://doi.org/10.1007/s11412-011-9123-y>
- Tegos, S., Demetriadis, S., Papadopoulos, P. M., & Weinberger, A. (2016). Conversational agents for academically productive talk: A comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, 11(4), 417–440. <https://doi.org/10.1007/s11412-016-9246-2>

- Tegos, S., Demetriadis, S., Psathas, G., & Tsiatsos, T. (2019). A configurable agent to advance peers' productive dialogue in MOOCs. *Lecture Notes in Computer Science*. [https://doi.org/10.1007/978-3-030-39540-7\\_17](https://doi.org/10.1007/978-3-030-39540-7_17)
- Terras, M., & Ramsay, J. (2015). Massive open online courses (MOOCs): Insights and challenges from a psychological perspective. *British Journal of Educational Technology*, 46(3), 472–487. <https://doi.org/10.1111/bjjet.12274>
- Webber, J., Scheuermann, B., McCall, C. M., & Coleman, M. S. (1993). Research on self-monitoring as a behavior management technique in special education classrooms. *Remedial and Special Education*, 14(2), 38–56. <https://doi.org/10.1177/074193259301400206>
- Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human-Computer Interaction*, 35(4–5), 356–373. <https://doi.org/10.1080/10447318.2018.1543084>
- Zhu, M., Sari, A. R., & Lee, M. M. (2020). A comprehensive systematic review of MOOC research: Research techniques, topics, and trends from 2009 to 2019. *Educational Technology Research and Development*, 68(4), 1685–1710. <https://doi.org/10.1007/s11423-020-09798-x>
- Zimmerman, B. J. (1998). Academic studing and the development of personal skill: A self-regulatory perspective. *Educational Psychologist*, 33(2–3), 73–86. <https://doi.org/10.1080/00461520.1998.9653292>
- Zimmerman, B. J. (2000). Attaining Self-Regulation. *Handbook of Self-Regulation*, 13–39. <https://doi.org/10.1016/b978-012109890-2/50031-7>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70. [https://doi.org/10.1207/s15430421tip4102\\_2](https://doi.org/10.1207/s15430421tip4102_2)

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Georgios Psathas<sup>1</sup>  · Stergios Tegos<sup>1</sup> · Stavros N. Demetriadis<sup>1</sup> · Thrasylvoulos Tsiatsos<sup>1</sup>

✉ Georgios Psathas  
gpsathas@csd.auth.gr

Stergios Tegos  
stegos@csd.auth.gr

Stavros N. Demetriadis  
sdemetri@csd.auth.gr

Thrasylvoulos Tsiatsos  
tsiatsos@csd.auth.gr

<sup>1</sup> Aristotle University of Thessaloniki, Thessaloniki, Greece