International Journal of Engineering Education Vol. 34, No. 3, pp. 865–878, 2018
Printed in Great Britain

0949-149X/91 $3.00+0.00

Uncovering Flipped-Classroom Problems at an Engineering Course on Systems Architecture Through Data-Driven Learning Design*

IRIA ESTÉVEZ-AYRES[1], JESÚS ARIAS FISTEUS[1], LUCÍA UGUINA-GADELLA[1,2], CARLOS ALARIO-HOYOS[1], and CARLOS DELGADO-KLOOS[1]

[1] Dpto. Ingeniería Telemática, Universidad Carlos III de Madrid, Avda. de la Universidad, 30, 29811, Leganés, Madrid, Spain.
E-mail: ayres@it.uc3m.es, jaf@it.uc3m.es, lucia.uguina@imdea.org, calario@it.uc3m.es, cdk@it.uc3m.es

[2] IMDEA Networks Institute, Av. del Mar Mediterráneo 22, 28918 Leganés, Madrid, Spain.

Flipped classroom is a student-centered methodology that can help engineering students to acquire the cross-curricular skills demanded by society. However, its effectiveness relies on the commitment of both instructors and students. In particular, this strategy requires students to work on a number of proposed activities before face-to-face classes. Then, in order to follow the most appropriate path in those classes, instructors need a reliable way to know at which degree their students worked on those proposed activities, what issues they encountered while doing them and which concepts need to be reinforced in class. This paper presents a case study of a flipped-classroom undergraduate engineering course. By using data-driven learning design and learning analytics techniques we show that: (1) by delaying their work on the course activities our students actually drove the course towards the traditional approach; (2) despite directly asking students at the beginning of a face-to-face class might seem to be an appropriate way of getting reliable information about their previous work, it may lead instructors to erroneous conclusions; (3) our students were strongly mark- and deadline-oriented, but even a small grade encouraged them to work on the assignments; (4) the gathering and checking of students' learning data before the class can help instructors to tailor the lesson design; and (5) if students did not work on pre-class activities, dedicating a small amount of time of the in-class lesson to explain the most difficult concepts can help students to be more efficient with their work, at the cost of losing some of the spirit of the flipped classroom.

Keywords: flipped learning; engineering education; learning analytics; learning design; challenges

1. Introduction

During the last decades, engineering education has evolved to meet society needs. Students should acquire professional engineering skills and, at the same time, cross-curricular skills, such as leadership, teamwork and self-learning [1–3]. The application of active learning and project-based approaches can help students to meet those demands, as these approaches promote most of the aforementioned cross-curricular skills [4, 5]. Moreover, when compared with traditional lectures, different studies agree that they are comparable when promoting the mastery of content, but active learning strategies outstand regarding the development of critical thinking and writing skills [6]. However, active learning courses are very sensitive to differences in students' backgrounds and learning paces, and the instructor should be provided with meaningful data in order to be able to react properly when problems arise [7].

The flipped classroom shares with other active learning strategies its core elements, i.e., it enhances traditional lectures by introducing practical activities in order to promote students' engagement [8]. However, the flipped classroom differs from other active learning approaches in where the meaningful learning takes place by shifting the workload of students inside and outside the classroom [9]. Instructors provide students with materials well before the class, and these materials need to be revised and worked on by the students before going into the classroom. Then, once in class, students engage in practicing through problem solving, critical discussion and collaborative learning [10, 11]. This way, instead of worrying about covering the whole course syllabus, there is space in the classroom for deepening into the most relevant concepts [12], and promoting meaningful learning [13]. Students actually learn to do something with the acquired knowledge, instead of memorizing some concepts without really understanding them [14], and some works claim that students prefer a flipped-classroom over traditional lectures [15]. This is also true in the field of computer science education, where the flipped classroom shows a positive effect on students' performance, attitudes and engagement [16]. However, there are works that disagree on the effectiveness of the flipped-classroom to improve students' performance [17, 18]. The debate is still open and, therefore, more studies on the application of the flipped classroom in different learning contexts are

* Accepted 12 January 2018.

necessary. In particular, our work targets the field of engineering education.

For the flipped classroom to be successful, it needs, as any other student-centered approach, commitment and engagement from the students [15, 19, 20]. This is especially important for students new to the flipped classroom, who are usually adamant and come unprepared to class, thus being unable to participate in the active learning phase of the course [19, 21]. In fact, the implementation of a flipped classroom strategy entails different challenges also for institutions and teaching staff. First of all, the implementation of flipped classroom strategies requires educational institutions to provide better infrastructures, which in turn requires an economic investment from them [22]. Student activities should be carefully planned and prepared by the instructors, which leads to much higher instructors' workload before the course [21]. Moreover, instructors' workload does not diminish during the course, as tutoring is a key aspect of the approach: doubts from students need to be promptly answered in order to keep them engaged with the flipped classroom strategy [23].

It is also important to notice that while the rhythm, depth and breadth of the class is imposed by the instructor in a teacher-centered approach, in student-centered courses, such as those implementing a flipped-classroom, all these three variables should ideally be tailored according to the different necessities of the students. Thus, there is a need to provide instructors with prompt and accurate information on the progress and difficulties experienced by their students. Armed with such data, instructors can make informed decisions based on a thorough analysis, thus changing when required the design of individual lessons or even of the whole course [24].

However, most of the studies which apply a flipped classroom strategy are focused on the analysis of students' self-reported work, voluntary questionnaires and final performance [25–26], instead of on real data of students' actions. Moreover, although data-driven learning design, i.e. the design of the learning experience based on the analysis of the data gathered from digital learning systems, is not a new concept [27, 28], there is not much work on the impact of using it combined with a flipped classroom strategy [29–31]. In this context, we pose the following research questions:

- RQ1: What is the impact on the performance and engagement of the students of the instructor applying data-driven learning design to a flipped classroom strategy?
- RQ2: Does the type of activity have an impact on students' engagement when applying a flipped classroom strategy?

In order to answer these research questions, this paper presents a case study on an undergraduate programming course that follows a flipped classroom approach. Regarding the first research question, an experiment was carried out over three weeks, where students were expected to use a simulation tool as part of their pre-class activities. The knowledge obtained by gathering data from this simulation tool was used to tailor the face-to-face lessons of the experimental group, thus improving instructor awareness. Meanwhile, the only information that was available for the instructor in the face-to-face lessons of the control group was the questions raised by the students. Data about the work performed by the students on the different activities of the course was gathered and analyzed to answer the second research question.

This paper is structured as follows: section 2 presents the methodology, materials and data sources used to conduct this research, as well as the design of the experiment; section 3 presents the results of this research; section 4 discusses the results; and, finally, section 5 presents the conclusions and future work.

2. Methodology

2.1 Background of the course

The empirical research is supported by an undergraduate engineering course on Systems Architecture. This 15-week course was taught in the fall semester of the second year of the bachelor's degree in Telecommunication Technologies Engineering. A total of 85 students enrolled in the course, from which 17 dropped before the end (they stopped attending lectures and exams). Therefore, only 68 completed it. Following University policies, the course followed a continuous assessment system, and each week students attended a lecture and a laboratory session of 100 minutes each. Lectures were delivered in a single large group, whereas students attended laboratory sessions in two smaller groups.

Our study is focused on the first 9 weeks of the course, where a flipped classroom strategy was applied. The remaining 6 weeks followed a project-based strategy instead. During those first 9 weeks, students were required to work at home on several activities prior to face-to-face lectures and laboratories. In the case of face-to-face lectures, pre-class activities introduced the topics of the current week, covering theoretical concepts. In the case of face-to-face laboratory sessions, pre-class activities consisted of programming exercises (in C language) of varied difficulty levels with the aim to prepare students to develop two bigger programming projects later in the course. Students were expected to work in teams to solve these activities by themselves, with the support of the instructor in tutoring sessions or in the online course forum. In face-to-face sessions, instructors assumed that pre-class activities had been already completed by students at home.

As the use of self-regulated learning techniques was one of the main learning outcomes of the course [32], the design of the flipped-classroom strategy for this course included two phases:

- (Phase 1) At home, prior to the lesson: students worked on pre-class material, composed by theoretical material and programming exercises with varied difficulty levels.
- (Phase 2) At face-to-face lessons in the classroom: in lectures, the instructor answered questions at the beginning of the session, and after that, students solved problems in a collaborative way; in lab sessions, students worked autonomously on programming assignments.

Students were presented with different types of activities during those first 9 weeks of the course:

- Formative activities, aimed to help students to identify their strengths and weaknesses, and instructors to early identify where students are struggling. These formative activities were not graded, and students had to complete them before the corresponding class.
- Summative activities, which belong to one of the following two subtypes according to their weight in the final grade of the student:
  - Regular summative activities, aimed at evaluating students at the end of each learning unit. They were part of the continuous evaluation and consisted of tests and in-class activities. Their total weight in the final grade of a student was approximately 98%. They will simply be called summative activities in the rest of the paper.
  - Optional activities, characterized by their very small weight in the total grade of the course (the remaining 2% of the final grade). They were aimed to encourage students' work during the development of each learning unit in what otherwise would have been formative activities.

One of the main drawbacks of following a flipped classroom strategy was the fact that students did not apprehend the concepts covered by the pre-class activities, thus hindering the development of the second phase of the flipped classroom strategy. When asked about their problems, most students stated that they covered all the pre-class work and that they understood everything. This happened even when students reported through anonymous questionnaires. However, when the instructors asked about concepts or exercises, the results of most students showed that their answers about their progress were unreliable. This could be due to the fact that students were not used to being responsible for their own learning, or that they overestimated their own level of understanding [33]. Without other reliable sources of information about the progress of the students, instructors relied on their own intuition to decide the rhythm, depth and breadth of the lesson.

2.2 Specific technologies for this research

This section is devoted to describing the specific technologies developed and used to conduct this research. First of all, it describes the C-mulator simulation tool, intended to be used by students prior to the face-to-face sessions. Its logs were processed and sent to the instructors of the experimental group before each face-to-face lesson. Then, it presents a monitoring tool used to track the work done by students in programming assignments, both at home and at laboratory sessions.

2.2.1 C-mulator

C-mulator, a web application that simulates the execution of C code over a simplified Von Neumann machine, was developed and used as part of phase 1 of the flipped classroom approach. This web application is aimed at understanding the basics of the C programming language. The decision to develop this tool was based on the feedback from students of previous editions of the course, where also a flipped classroom strategy was followed. Since the first edition of the course (Fall 2009), one of the recurring topics on students' complaints was the difficulty of memory management in the C programming language, and the problems they faced because of having to lower the level of abstraction from the high-level Java language used in previous programming courses to the middle-level C language [32]. This difficulty is also pointed out by several studies in the literature [34–36]. Simulators are invaluable tools when giving insights of the internal behavior and interaction of the different units that constitute a computer architecture [37, 38], and their use as a supplementary tool to lectures can help students to defy their barriers in cognitive learning [39].

C-mulator works with an input C file, which must be uploaded by the student. The C file must contain the ''main'' function, and students can use any other function from the standard C library or define their own functions. Then, the student can run the C program step-by-step, visualizing at the same time the current state of the internal memory (heap and stack), the code that is being executed and the standard output produced by the program. Logs about the use of C-mulator in this course were gathered and processed, in order to obtain more information about the progress of the students.

2.2.2 *Other used technologies*

The programming exercises required a Linux environment and some specific software. Thus, students were provided at the beginning of the course with a Virtualbox virtual machine [40] that replicated the configuration of the computers used in laboratory sessions. This way, students could work at home without having to install the environment themselves. Programming exercises were structured in directories and delivered to students through the Subversion version control system [41]. Those directories were the de facto workspace for students.

A data gathering tool based on [42] and [43] was developed to collect CLI (Command Line Interface) events from students' workspaces. The tool tracked text editors (emacs and kate), compilers (gcc), debuggers and analysis tools (gdb and valgrind) and version control commands (svn). The virtual machine provided to the students had the data gathering tool installed and a script to turn the gathering process on and off. The main page of the course informed students about what was tracked and how to turn the data gathering tool off. Students were also provided with an installer to deploy the tracking system in other computers if they wished. The tracking was automatically disabled at the end of the semester.

2.3 *Design of the experiment*

In order to answer the first research question, "RQ1: What will be the impact of a data-driven learning design performed by the instructor when applying a flipped classroom strategy?" we conducted a three-week experiment during weeks 5, 6 and 7, where one of the groups of laboratory sessions was used as experimental group and the other one as control group. The experimental group had initially 41 students, from which 33 completed the course. The control group had 44 students from which 35 completed the course. Both groups shared theoretical lectures. Students were divided into 19 teams within the experimental group and 22 teams within the control one.

The timeline of the experiment was as follows:

- At the fifth week of the course, just before starting to use C-mulator, a pre-test was carried out.
- C-mulator was used during weeks 5, 6 and 7 as part of the formative activities students should prepare before face-to-face laboratory sessions.
- In the ninth week of the course a post-test was carried out.

Several simulations were prepared for the face-to-face lectures during the weeks of the experiment. In addition, and as part of phase 1 of the flipped classroom strategy, several additional simulations were prepared for each laboratory session as pre-class formative activities, meant to be used between reading the theoretical material and implementing the C programming exercises. Before the first pre-class activity involving the use of C-mulator by individual students, the instructor, at the theoretical lecture, explained through different examples its basic functionality. Both experimental and control groups were expected to complete the C-mulator pre-class activities. The deadline for completing all the formative activities was the start of the lesson itself, while the deadline for completing all the summative activities was one week after the lesson.

Instructors also prepared the design of the laboratory session as a learning graph with several learning paths, depending on the doubts and homework done by the students. However, the behavior of the instructor was different depending on the group:

- In the control group, the instructor assumed that, if no question was raised by them (in fact, none was actually raised), students had completed the pre-class activities and, therefore, covered the needed pre-class concepts.
- In the experimental group, and before each lab session, the instructor checked whether the students had used the simulation tool and with which simulations they had interacted. Using this information, the instructor prepared the session and decided on the most suitable learning path to take. More specifically, she dedicated some time at the beginning of the lab sessions to the concepts they should have learned in those activities at home. This entails a major improvement for the instructor in the redesign of the class, as this redesign is done based on actual data (experimental group), instead of on self-reported information (control group).

Answering the second research question, "RQ2 Does the type of activity have an impact on students' engagement when applying a flipped classroom strategy?", required conducting an anonymous and voluntary survey, gathering opinions from a focus group and collecting CLI events from the students during the first 9 weeks of the course, as explained in the next section.

2.4 *Data collection methods*

This experiment comprised different data sources with different timelines:

- CLI events (used for answering RQ1 and RQ2): collected during the first nine weeks of the course by the tracking tool.

- C-mulator logs (RQ1): gathered during the 3-week C-mulator experiment.
- Quantitative tests (RQ1), in the form of a pre-test before the C-mulator experiment and a post-test after it. Both pre-test and post-test consisted of a non-graded questionnaire about the contents of the course up to that class, focusing especially on practical questions related to the lab assignments. The post-test assessed the concepts studied during the C-mulator experiment.
- An anonymous and voluntary self-reported survey at the end of the C-mulator experiment (RQ2), composed of two open questions about the course (one about the most positive aspect of the course and another one about the most negative aspect), a 5-point Likert-type question about the utility of C-mulator, and a yes-no question about whether students recommended its use in future editions of the course. The aim of this survey was to evaluate the students' general perception about the flipped classroom methodology and tools used.
- Focus group opinions (RQ2): at the end of the course, several meetings were held with a focus group composed of volunteers that had passed the course with varied degrees of performance. The objective of the focus group was to gather general opinions about the course, focusing the discussion on the students' perception of usefulness of the different types of activities.

2.5 Data analysis methods

As data of different types were gathered in the experiment, different quantitative and qualitative analysis methods had to be used:

- Qualitative data: they were obtained from the opinions gathered from students in the anonymous self-report survey and in the focus group. The opinions from the survey were analyzed and then classified per topic by the instructors following the mixed method proposed in [44].
- Quantitative data from C-mulator logs: the number of students that used the tool, when they used it and which programs they used were analyzed.
- Quantitative data from Students' Grades: the grades from the experimental and control groups were analyzed using descriptive statistics (mean and standard deviation). Then, a t-test analysis was conducted to determine if the two groups differed significantly. This analysis was performed for the grades of the pre-test and the post-test.
- Quantitative data from Students' Workspaces (CLI events): these data were used to perform different analysis regarding the number of activities students attempted, the invested time per activity, and when the students started each activity. Given the complexity of these data, a more detailed explanation about it follows below.

Each CLI event collected from the workspace of the students contained the following information: (1) a timestamp with the instant in which the action happened; (2) the learners' identity; (3) the identity of the environment in which the event was created, e.g., the identifier of a specific virtual machine instance or laboratory computer; (4) the event type (gcc, svn, etc.); (5) the current directory in which the action was executed (the value of the Linux PWD variable); (6) the whole command entered by the student; (7) the finishing status of the command; and (8) the standard output and standard error generated by the command.

After gathering the events, as each directory of the workspace corresponds to a specific laboratory and exercise, each event was associated to a laboratory session, exercise and type of activity (formative, optional or summative). As students worked in teams, the events were also annotated with the specific team and group (experimental or control). After that, events were grouped in working sessions, assuming that a working session begins when the student generates an event and finishes when there is no activity from that student for at least one hour. Finally, as the deadline for each summative and formative activity was known, events were also annotated with the difference between their timestamp and the deadline of their corresponding activity, a negative value meaning that the action was performed before the deadline, and a positive value meaning that the action was performed after the deadline.

Almost all the analyses regarding the work done by students in programming assignments refer to the work performed by teams instead of individual students, as it is difficult to distinguish individual work in laboratory assignments since students frequently sit together in pairs in front of a single computer.

The first analysis focused on the percentage of attempted activities by each team, i.e., the percentage of activities that had at least one associated event. Results were computed separately for each activity type (formative, optional and summative). Descriptive statistics were used to analyze data separately from the experimental and from the control group. In this case, as the data was not normally distributed, the central tendency of the data was analyzed using the median (M_D), and its variability using the lower (Q1) and upper (Q3) quartiles. Additionally, the correlation between the individual post-test scores and the percentage of attempted activities by activity type was also computed.

The second analysis focused on the time invested per exercise by activity type. In order to obtain it, the different identified student working sessions were grouped in teams and then processed to obtain the periods of time devoted to only one specific exercise. Then, the total amount of time per exercise per team was computed. These figures were grouped by activity type and, then, analyzed. In this case, the data followed a normal distribution. Thus, the mean and standard deviation were used as descriptive statistics to separately analyze data from the experimental and control groups. After this, several t-tests were performed to determine whether the experimental group behaved differently from the control group. Then, the total amount of invested time per activity type by each team was computed and descriptive statistics for non-normal distributed variables were used to analyze data separately from the experimental and from the control group (median, lower and upper quartiles). The influence on the individual post-test scores of the total amount of invested time per type of activity was also assessed using the Pearson correlation.

The third analysis focused on the time, relative to the activity deadline, at which students started working on each activity. All the gathered events of each team were grouped by exercise and the instant of the first command was recorded. Then, the exercises were grouped by activity type and an analysis based on descriptive statistics (mean and standard deviation, as they followed a normal distribution) was conducted for both experimental and control groups, and several t-tests were performed to determine whether the experimental group behaved differently from the control group.

Finally, parts of these analyses were repeated for CLI events gathered during the first nine weeks of the course, in order to study the long-term behavior of the students regarding the application of the flipped classroom.

3. Results

3.1 Research question 1

Only 4 out of 33 students of the experimental group used C-mulator as part of their pre-class activities. Thus, this section aims to unveil whether the approach taken by the instructor (investing a little time, approx. 30% of the lesson to explain pre-class concepts in the experimental group) had an impact on the performance and engagement of the students.

The results of the pre-test are shown in Table 1. There was no significant difference in the scores between the experimental and the control groups (p-value = 0.55). The results of the post-test are also shown in the Table 1. The experimental group scores (N = 33, M = 4.14, SD = 3.49, N being the number of students) outperformed the control group scores (N = 35, M = 3.04, SD = 3.08). However, the difference between both groups is not significant (p-value = 0.08).

The total number of CLI events collected by the tracking tool during the three weeks of the experiment was 27,298 (10,372 from the experimental group and 16,926 from the control group). There were no optional activities during these three weeks, i.e., students were expected to work only on formative (all of them pre-class) and summative (all of them in-class) activities.

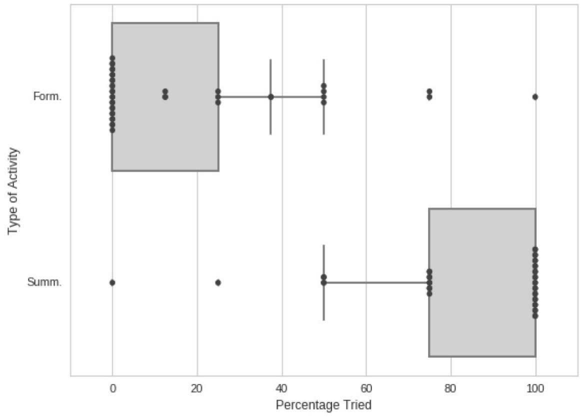
The behavior of both the experimental and control groups was very similar regarding how many formative and summative activities were attempted by each team, as shown in Fig. 1. In fact, the median, first quartile and third quartile were exactly the same for both groups (Summative activities: Q1 = 75%, M_D = 100%, Q3 = 100%; Formative activities: Q1 = 0%, M_D = 0%, Q3 = 25%, with N_{Teams} = 22 in the experimental group, and N_{Teams} = 19 in the control group).

When analyzing the impact on the individual post-test score of the percentage of attempted activities depending on the type of activity, there is a significant difference between the experimental and the control group, as shown in Table 2. As expected, the individual post-test scores showed a high and significant correlation in both groups with the percentage of attempted summative activities. However, the relationship with the number of attempted formative activities was different in the two groups; while the grades of the experimental group showed no significant correlation (N = 33, r = 0.1213, p = 0.5011, with N the number of students), the grades of the control group showed a high and

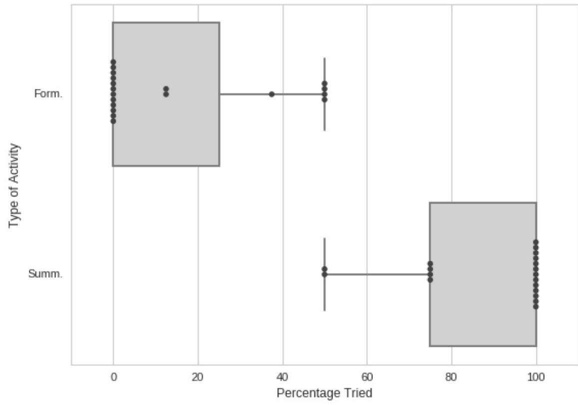
Table 1. Comparison of Pre test and Post Test Grades (t-test one-tail)

	Experimental Group			Control Group						
Test	**N_{Teams}**	**M**	**SD**	**N_{Teams}**	**M**	**SD**	**t**	**df**	**p**	**95% Confidence Interval**
Pre-Test	33	4	3.56	35	4.12	3.48	–0.1369	65.5	0.55	(–1.54, Inf)
Post-Test	33	4.14	3.49	35	3.04	3.08	1.3756	63.8	0.08	(–0.23, Inf)

Note: N_{Teams} = size of the sample (number of teams), M = Mean, SD = Standard Deviation, t = t-test, df = Degrees of freedom.



(a) Experimental Group



(b) Control Group

Fig. 1. Percentage of Attempted Activities (during the C-mulator experience).

Table 2. Correlation Analysis between the Post-Test grades and the Percentage of Attempted Activities

	Experimental Group				**Control Group**			
Type of Activity	**N**	**r**	**r^2**	**p**	**N**	**r**	**r^2**	**p**
Summative	33	0.4606	0.2121	0.007	35	0.5476	0.2999	0.0007
Formative	33	0.1213	0.0147	0.5011	35	0.4405	0.1941	0.0081

Note: N = Number of Students, r = Pearson Correlation, r^2 = coefficient of determination, p = significance of the correlation.

Table 3. Comparison of Experimental and Control groups regarding the Invested Time (in minutes) per activity by type of activity (t-test one-tail)

	Experimental Group			**Control Group**						
Type of Activity	**N_T**	**M**	**SD**	**N_T**	**M**	**SD**	**t**	**df**	**p**	**95% Confidence Interval**
Summative	76	212.7	228.5	88	279.7	283.9	–1.6634	161.2	0.049	(–Inf, –0.36)
Formative	152	5.9	22.3	176	9.7	31.3	–1.267	314.9	0.103	(–Inf, 1.14)

Note: N_T = size of the sample (total number of activities), M = Mean, SD = Standard Deviation, t = t-test, df = Degrees of freedom.

significant correlation (N = 35, r = 0.5011, p = 0.0081).

Their behavior was also different regarding the amount of time invested per activity by type of activity (see Table 3 and Fig. 2). Although there is no significant difference in the amount of time invested in each of the formative activities (p-value = 0.103), which makes sense due to the small number of groups that attempted the formative activities, the control group invested one hour more per summative activity on average, being this difference significant (p-value = 0.049).

An analysis of the total amount of time invested per type of activity was also performed (see Fig. 3). Regarding the total time invested in the formative activities, the behavior was very similar between the groups (Q1 = M_D = 0%, Q3 = 0.935 hours, with N_{Teams} = 22 in the experimental group, and Q1 = M_D = 0%, Q3 =1.341 hours, N_{Teams} = 19 in the control group). However, regarding the total time invested in the summative activities, the variability was lower in the experimental group (Q1= 7.889 hours, M_D = 10.137 hours, Q3 = 15.949 hours, with N_{Teams} = 22) than in the control group (Q1= 7.82 hours, M_D = 20.742 hours, Q3 = 23.907 hours, with

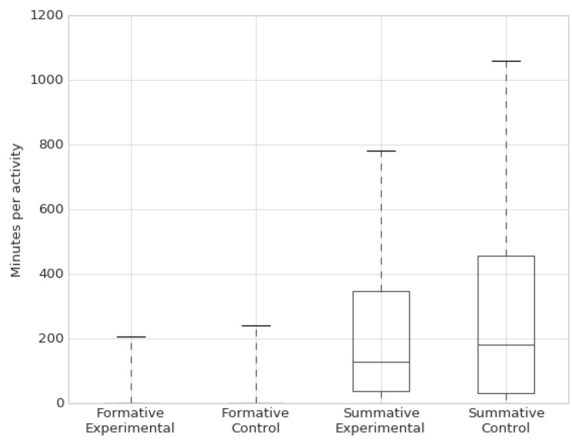
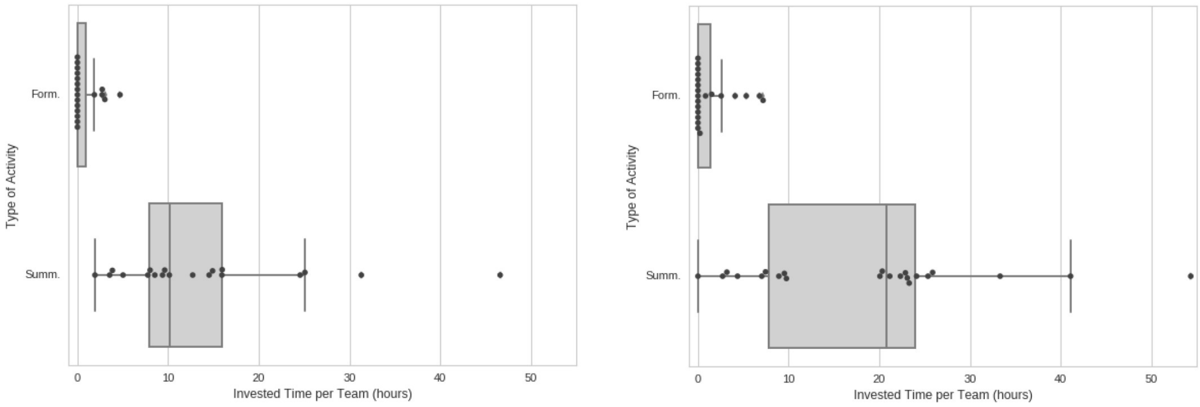


Fig. 2. Average time in minutes per Activity by Type of Activity (during the C-mulator experience).



(a) Experimental Group (b) Control Group

Fig. 3. Total time invested per type of Activity (during the C-mulator experience).

Table 4. Correlation Analysis between the Post-Test grades and the Total Amount of Invested Time per Type of Activity

	Experimental Group				Control Group			
Type of Activity	**N**	**r**	**r^2**	**p**	**N**	**r**	**r^2**	**p**
Summative	33	–0.0916	0.0084	0.5739	35	0.5224	0.2729	0.0003
Formative	33	0.0384	0.0014	0.8140	35	0.5683	0.3230	$5{\cdot}10^{-5}$

Note: N = Number of Students, r = Pearson Correlation, r^2 = coefficient of determination, p = significance of the correlation.

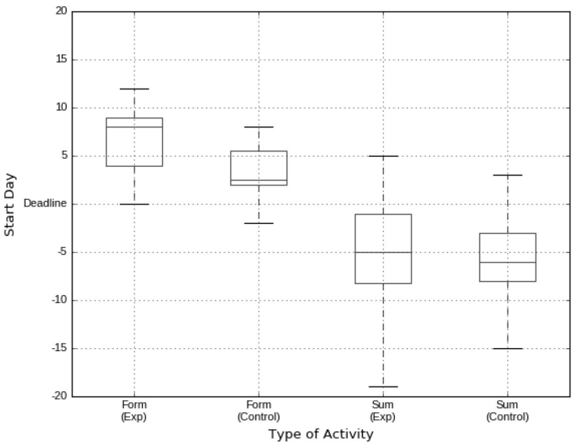


Fig. 4. Starting Day by Type of Activity (during the C-mulator experience)

N_{Teams} = 22). The impact of this variable on the individual post-test score was analyzed (see Table 4). Surprisingly, in the case of the experimental group the amount of time dedicated to both the summative and formative activities shows a very low and non significant correlation with the post-test score, while these same variables exhibit a high and significant correlation in the case of the control group (p-value = 0.003 for the correlation between the total amount time invested on summative activities and the post-test score, and p-value = $5 \cdot 10^{-5}$ in the case of formative activities).

As the flipped classroom relies on the work performed by students prior to the lesson, the instant at which they started working on the different assignments is also informative. As shown in Fig. 4 and Table 5, the control and experimental groups behaved similarly regarding to when teams started working on the summative activities, the difference between the groups was not significant (p-value = 0.98). However, the behavior regarding when teams started working on the formative activities was different (the control group started on average four days before the experimental group), this difference being significant (p-value = 0.02).

Table 5. Comparison of Experimental and Control groups regarding the Starting day of Activity by type of activity (t-test two-tail)

	Experimental Group			Control Group						
Type of Activity	**N_A**	**M**	**SD**	**N_A**	**M**	**SD**	**t**	**df**	**p**	**95% Confidence Interval**
Summative	68	–5.6	7.9	72	–5.6	4.37	0.0216	103.3	0.98	(–2.14, 2.19)
Formative	21	7.3	6.6	30	3.4	2.61	2.5268	24.3	0.02	(0.72, 7.14)

Note: N_A = size of the sample (total number of attempted activities), M = Mean, SD = Standard Deviation, t = t-test, df = Degrees of freedom.

3.2 Research question 2

This section presents the results of the analysis of students' activity both in the laboratories and in their virtual machines during the first nine weeks. The total number of gathered events during them was 80,259. Additionally, it presents the results of the anonymous survey and the focus group.

During the weeks that were not part of the C-mulator experiment (i.e. weeks 1–4 and 8–9), there were only optional activities (both pre-class and in-class activities). Thus, this section mainly focuses on optional summative activities, comparing them to the behavior of the students regarding formative and regular summative activities. It is interesting to notice that, although the additional weight in the final score of the optional activities was quite small (only 2%), the median of the percentage of attempted activities was higher than 75% (N_{Teams} = 41, Q1 = 68.4%, M_D = 78.9%, Q3 = 94.7%). Regarding the time invested per activity by activity type (see Fig. 5), as said before, students spent more time on the summative activities, and invested almost no time on the formative activities. The time invested per optional activity was small compared to summative activities, approximately half the time (optional activities: N_T = 779, M = 97.3 minutes, SD = 130.3 minutes; summative activities: N_T = 164, M = 248.7 minutes, SD = 261.85 minutes, with N_T the total number of activities). However, as optional activities were conceived to engage students with what would have otherwise been formative activities, their level of difficulty was lower than the level of difficulty of the summative activities and, then, less time was needed to complete them.

Fig. 6 shows when students started working on the activities depending of the type of activity. Activities are divided between *pre-class* activities, which students should try (and complete, if possible) before going to class and *in-class* activities, which students should start in the classroom. The deadline of the pre-class activities was the start of the lesson itself, while the deadline of the in-class activities was their submission deadline. It can be seen that, regarding in-class activities (optional and summative activities), students started working before the deadline (day 0) but after the face-to face session (day –7). In average, summative activities were started one day or two after the face-to-face session (N_A = 139, M = –5.6 days, SD = 6.3 days, being N_A the total number of attempted activities). The same happened with the in-class optional activities: they were started after the face-to-face session and before their deadline, although in this case students waited until four days before the deadline to start them (N_A = 467, M = –3.9 days, SD = 4.7 days). What happens with the (graded) pre-class optional activities is interesting: students were advised to complete them before the face-to-face session but, in average, they started them during the face-to-face session (N_A = 142, M = –6.8 days, SD = 3.4 days). Regarding formative activities, students started working on them long after their deadline (N_A = 51, M = 5 days, SD = 5.1 days).

In order to understand the difference in the level of engagement of the students depending on the type of activity, an anonymous survey and a focus group were used. The anonymous self-reported survey was answered by 17 students. None of the students that answered the survey had used C-mulator. The reasons why they had not used it varied from "high workload" to "having understood everything when explained in class". However, all of them agreed on the usefulness of the web application (all of them had seen the application working during the instructor explanation at the theoretical lesson previous to the C-mulator experiment) and, when asked about whether the application should be wholly deployed for the next edition of the course, 55.6% agreed, while 44.4% answered "I don't know". This questionnaire contained an open question about the most positive aspect of

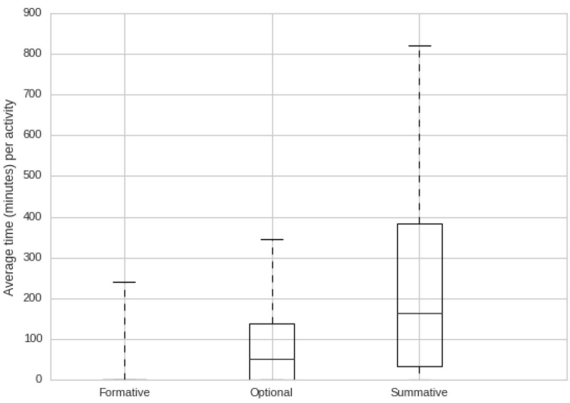


Fig. 5. Average time in minutes per Activity (by Type of Activity).

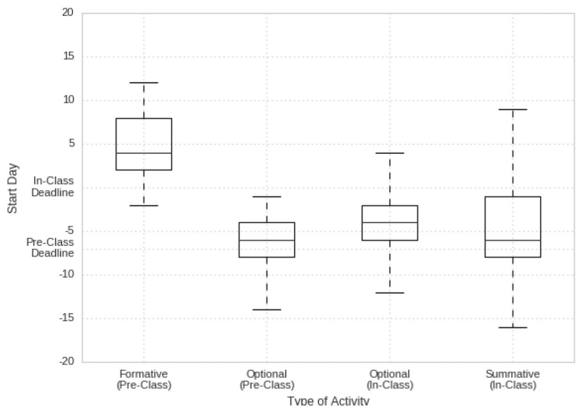


Fig. 6. Starting Day by Type of Activity.

Table 6 Anonymous Survey Comments by Category

Category	N	Positive (%)	Negative (%)
Workload	18	16.6%	83.3%
General Methodology	8	62.5%	37.5%
Theoretical Sessions	3	66.6%	33.3%
Lab Sessions	1	100%	0%
Course Changes	2	50%	50%
Teaching Support	6	83.3%	16.6%
Collaborative Learning	3	66.6%	33.3%
Evaluation	1	0%	100%
C-mulator	1	100%	0%
Total	43	46.5%	53.5%

Note: N = number of comments.

the course and another one about the most negative aspect. Instructors analyzed these comments independently and classified them into positive and negative regarding the category. As shown in Table 6, 18 out of the 43 identified comments were related to the course workload, being the 83.3% of these comments negative, and mainly referring to students' complaints about the amount of work needed to complete the course.

The focus group involved 16 volunteers that had passed the course. During the focus group, all the participants agreed on the extremely "mark-oriented" behavior of their classmates. Their shared opinion was that their classmates dismissed the utility of all the formative activities that did not have a direct impact on their final score. They also pointed out that formative activities were generally performed and reviewed before the tests instead of before the lessons they were intended for. From their shared point of view, only summative activities were done in general before their deadline.

4. Discussion

The results presented in the previous section allow us to answer the two research questions (RQ1 and RQ2) and to propose some solutions to the underlying problems of a flipped classroom approach that can be useful in other similar environments.

Concerning RQ1 ("What is the impact on the performance and engagement of the students of the instructor applying data-driven learning design to a flipped classroom strategy?"), results from the C-mulator experiment can be discouraging at a first glance: both groups seem to be very similar if only the performance on the pre-test and post-test is analyzed. Even the number of attempted activities (formative and summative) and the amount of time invested solving formative activities are very similar in both groups. However, the post-test scores of the control group show a high and significant correlation with the number of attempted formative activities and the total amount of time invested on solving summative and formative activities. These results, along with the fact that there is no significant difference in their post-test scores, could mean that the redesign of the lecture done by the instructor of the experimental group by taking into account the data coming from C-mulator compensated the fact that some teams did not work on the formative activities. Remember that, being aware of the little pre-class work done by the students in the experimental group, the instructor of the experimental group devoted 30 out of 100 minutes of the laboratory session to cover the most relevant concepts from the pre-class material. On the contrary, the instructor of the control group, as students did not raise doubts, devoted no time to cover concepts from the pre-class material. In addition, the fact that the students in the experimental group invested less time on the summative activities than the students in the control group while obtaining similar scores in the post-test might suggest that those 30 minutes of instructor's explanations, despite only covering some of the concepts of the pre-class activities, helped them to use their working time more efficiently than the control group.

Thus, we believe that, considering that this is the first course with a flipped classroom approach for the majority of the students, the use of a data-driven design in the lectures could be beneficial to guide the students through this new approach and to help them to be more efficient with their working time.

Concerning RQ2 ("Does the type of activity have an impact on students' engagement when applying a flipped classroom strategy?"), we found out that, in this specific context, students do not follow the suggested schedule (pre-class activities at home, in-class activities at face-to-face lessons). Instead of that, they tuned the schedule into a more traditional one (pre-class activities at face-to-face lessons, in-class activities at home, after the lesson), in both the experimental and control groups. Pre-class activities were designed to ease the understanding of the in-class activities, which were intended to be discussed and solved during the face-to-face session, as they contained more complex topics, thus devoting face-to-face time to more meaningful learning. However, when these pre-class activities were not graded (formative activities), most students did not even attempt them, and those who attempted them started working on them after the face-to-face lesson and spent only a small amount of time on each one. This finding aligns with the opinion of the focus group regarding the behavior of their classmates as extremely mark-oriented. It also aligns with the low use of C-mulator, which being a formative activity (with no associated mark) was considered "useful but not worth the time, being not graded" by most students.

It is interesting what happened with the pre-class graded optional activities, which students were advised to complete before the face-to-face lesson: students started these activities during the lesson itself, thus losing their pedagogic intent and delaying the start of the in-class activities. On the positive side, when an activity was graded, students worked hard, no matter the weight of the activity in the final score, even on optional activities, whose aggregated weight in the final grade was just a 2%.

This finding aligns with the lack of receptivity of students to the structure of the course, which was reported in [16] as one of the challenges when applying a flipped classroom strategy. We believe that, although not all the pre-class activities should be graded, a good mixture of graded and non-graded activities could be beneficial, especially if the graded activities are designed to rely on the successful completion of the non-graded ones, to ensure that students follow the learning schedule. The weight of those pre-class graded activities is up to the instructor.

Being their first flipped classroom experience, the heavy workload of the course was expected to be the main complaint from the students in the anonymous survey at the end of the C-mulator experiment. In fact, in previous editions of this course, the number of negative opinions about the workload showed a peak at the middle of the semester (the time at which this anonymous survey was made), which decreased towards the end of the course, increasing at the same time positive comments from the students about their perceived learning [32]. This is a common comment from students in flipped classroom courses, as students regard the shift of workload from post-class or time before the final exam to pre-class as "extra work", not being able to acknowledge the pedagogic value of the pre-class work [43, 44]. We believe that the decrease on the complaints about the workload towards the end of the course shows that, when given enough time to try the approach, students eventually acknowledge its pedagogic value.

This study presents some limitations which are worth mentioning. The success of the flipped classroom depends strongly on the commitment of the students. This experience aimed at uncovering some of the problems that instructors may face when adopting a flipped classroom strategy. During the C-mulator experiment, the main difference between the experimental and the control groups was the awareness of the instructor. In the control group, the instructor assumed that the pre-class activities had at least been attempted by the students and moved forward towards more difficult concepts. In the experimental group the instructor, after having checked that almost no work had been done by the students, devoted part of the class to briefly explain the most important concepts they should had learned at home. This works to an extent, as it loses part of the flipped classroom spirit. However, it allows us to compare the behavior of a "pure" flipped classroom group (the control group) with a mixed flipped classroom group (the experimental group). Another limitation is that the experiment only lasted three weeks due to context constraints, since after the first nine weeks the whole laboratory time is devoted to project-based learning, with the implementation of a more complex software project. However, a longer experiment would have been useful to assess the implication of using a data-driven learning design during a whole course.

As said before, this is the first course in the curriculum of the students following this type of strategy, and thus some reluctance from students was expected [19, 21]. In anticipation to this problem, the rationale behind this pedagogical approach was explained during the first lecture of the course, to try to augment the engagement of students, as recommended in [47]. Therefore, several introductory readings about flipped classroom [10], active learning [8, 46] and project-based learning [49] were recommended to students. Moreover, all the material of the course and its schedule (including the deadline for each formative, summative and optional activity) was available to students from the first day of the course. Despite this, most students did not commit with the flipped classroom strategy. It would be interesting to repeat this experiment on the same students but further on the curriculum (in upper courses), in order to assess whether their maturity and previous experiences with courses that follow similar strategies could be factors to achieve a successful flipped classroom, as pointed out by [50]. Furthermore, although this was the first course in their curriculum that followed this strategy, this is a second-year course that relies on previous background provided by two first-year programming courses. We found out that many students enrolled in this course despite not having passed the two previous programming courses (the University policies allow them to do so). Students with an inadequate background face more difficulties to follow the flipped classroom strategy [51], hindering the success of the experiment.

5. Conclusions and future work

The flipped classroom can provide students with a path towards a more meaningful learning. In the field of computer science education, a positive effect of the flipped classroom on students' performance, attitudes and engagement is reported in the literature. However, student commitment is required to achieve this positive effect, and instructors need accurate information about learners' attitudes and prior work in order to apply data-driven learning design to the teaching sessions, thus being able to achieve an environment where the learning depth, rhythm and breadth of the lessons are determined by students.

In this work, we have found that students enrolled for the first time in a flipped-classroom course are more than reluctant to follow this approach. Even if the benefits of this approach are explained and motivated in detail, the use of extrinsic motivation, such as scoring the activities (even with very small weights), is needed to get students to work. Fortunately, it seems that the weight of those activities in the final score is not very important: students in this course worked hard and invested time on graded activities regardless their weight. On the contrary, they dedicated almost no time to non-graded formative activities, which did not contribute to their final score. Apart from being mark-oriented, students also showed to be deadline-oriented. Despite the instructor suggesting doing some activities before face-to-face sessions, because their actual submission deadline was one week after, most students delayed working on them until the face-to-face session. The awareness of the instructor about this situation (and her explanations about the more difficult concepts covered by the pre-class material) seems to help students to be more efficient with their work. Although these results cannot be extrapolated to other learning situations with students with more self-regulated learning skills, and despite the fact that more research is needed in this area, the findings of this research aligned with the instructor's intuition about the behavior of her students and with other previous studies.

Future lines of work include studying the effect of fostering students' motivation in the course to better align with the flipped classroom paradigm without having to grade every single activity. Thus, for the next editions of the course we are mixing graded and non-graded activities on the schedule to try to ensure the completion of the pre-class activities before class. We are also working on the relationship between students' self-regulated learning skills, their performance in a flipped-classroom course and the possible improvement of these skills after this kind of courses. Moreover, during this academic year, a first-year programming course of the same degree has followed for the first time a flipped classroom strategy by reusing some MOOCs (Massive Open Online Courses) developed by the instructors themselves. Therefore, the students of our next edition of the course will be more familiar with the paradigm, allowing us to assess the impact of having students experienced in the flipped classroom.

Acknowledgments—This work was partially funded by: the Madrid Regional Government (Comunidad de Madrid), through the *eMadrid Network* (S2013/ICE-2715); by the Spanish Ministry of Competitiveness and Economy, through projects RESET (TIN2014-53199-C3-1-R) and *AUDACity* (TIN2016-77158-C4-1-R) and through the thematic network of excellence, *SNOLA* (TIN2015-71669-REDT); and by the European Commission, through Erasmus+ projects *MOOC-Maker* (561533-EPP-1-2015-1-ESEPPKA2-CBHE-JP), *SHEILA* (562080-EPP-1-2015-1-BEEPPKA3-PI-FORWARD), *COMPASS* (2015-1-EL01-KA203-014033), and *COMPETEN-SEA* (574212-EPP-1-2016-1-NL-EPPKA2-CBHE-JP).

References

1. T. Litzinger, L. R. Lattuca, R. Hadgraft and W. Newstetter, Engineering education and the development of expertise, *Journal of Engineering Education*, **100**(1), 2011, pp. 123–150.
2. S. Sheppard, K. Macatangay, A. Colby and W. M. Sullivan, *Educating engineers: Designing for the future of the field*, Jossey-Bass Inc Pub, San Francisco, USA, 2008.
3. B. Trilling and C. Fadel, *21st century skills: Learning for life in our times*, Jossey-Bass Inc Pub, San Francisco, USA, 2009.
4. A. Martínez-Monés, E. Gómez-Sánchez, Y. A. Dimitriadis, I. M. Jorrín-Abellán and B. Rubia-Avi, Multiple case studies to enhance project-based learning in a computer architecture course, *IEEE Transactions on Education*, **48**(3), 2005, pp. 482–489.
5. I. C. Moura and N. van Hattum-Janssen, Teaching a CS introductory course: An active approach, *Computers & Education*, **56**(2), 2011, pp. 475–483.
6. C. C. Bonwell and J. A. Eison, *Active Learning: Creating Excitement in the Classroom, 1991 ASHE-ERIC Higher Education Reports*, ERIC Clearinghouse on Higher Education, The George Washington University, Washington, DC, USA, 1991.
7. C. Bouton and R. Y. Garth, Students in learning groups: Active learning through conversation, *New Directions for Teaching and Learning*, **1983**(14), 1983, pp. 73–82.
8. M. Prince, Does active learning work? A review of the research, *Journal of Engineering Education*, **93**(3), 2004, pp. 223–231.
9. J. Bergmann and A. Sams, *Flip your classroom: Reach every student in every class every day*, International Society for Technology in Education, 2012.
10. B. Tucker, The flipped classroom, *Education Next*, **22**(1), 2012, pp. 82–83.
11. K. L. Bishop and M. A. Verleger, The flipped classroom: A survey of the research, *ASEE National Conference Proceedings*, Atlanta, USA, June 23–26, 2013, pp 1–18.
12. M. Vaughan, Flipping the learning: An investigation into the use of the flipped classroom model in an introductory teaching course, *Education Research and Perspectives*, **41**, 2014, pp. 25–41.
13. M. D. Estes, R. Ingram and J. C. Liu, A review of flipped classroom research, practice, and technologies, International HETL (Higher Education Teaching and Learning Association) Review, **4**, article 7, 2014.
14. J. Michael, In pursuit of meaningful learning, *Advances in Physiology Education*, **25**(3), 2001, pp. 145–158.
15. T. Lucke, Using learning analytics to evaluate the effectiveness of the flipped classroom approach, *25th Annual Conference of the Australasian Association for Engineering Education: Engineering the Knowledge Economy: Collaboration, Engagement & Employability*, 2014, pp. 1156–1164.
16. M. N. Giannakos, J. Krogstie and N. Chrisochoides, Reviewing the flipped classroom research: reflections for computer science education, *Proceedings of the Computer Science Education Research Conference*, Atlanta, USA, March 5–8, 2014, pp. 23–29.
17. D. González-Gómez, J. S. Jeong and D. A. Rodríguez,

Performance and perception in the flipped learning model: an initial approach to evaluate the effectiveness of a new teaching methodology in a general science classroom, *Journal of Science Education and Technology*, **25**(3), 2016, pp. 450–459.
18. S. B. Velegol, S. E. Zappe and E. Mahoney, The evolution of a flipped classroom: Evidence-based recommendations, *Advances in Engineering Education*, **4**(3), 2015, pp.1–37.
19. L. Chellapan and J. van der Meer, Challenges in Implementing the Flipped Classroom Model in Higher Education, in *Handbook of Research on Active Learning and the Flipped Classroom Model in the Digital Age*, IGI Global, 2016, pp. 352–365.
20. M. Cole, Using Wiki technology to support student engagement: Lessons from the trenches, *Computers & education*, **52**(1), 2009, pp. 141–146.
21. C. F. Herreid and N. A. Schiller, Case studies and the flipped classroom, *Journal of College Science Teaching*, **42**(5), 2013, pp. 62–66.
22. C. K. Lo and K. F. Hew, A critical review of flipped classroom challenges in K-12 education: possible solutions and recommendations for future research, *Research and Practice in Technology Enhanced Learning*, **12**(1), 2017, p. 4.
23. J. Zhang, Y. Wang and B. Zhang, Introducing a new teaching model: Flipped classroom, *Journal of Distance Education*, **4**, 2012, pp. 46–51.
24. P. S. Steif and A. Dollár, Study of Usage Patterns and Learning Gains in a Web-based Interactive Static Course, *Journal of Engineering Education*, **98**(4), 2009, pp. 321–333.
25. C. Alario-Hoyos, I. Estévez-Ayres, C. D. Kloos and J. Villena-Román, From MOOCs to SPOCs . . . and from SPOCs to Flipped Classroom, *Proceedings of the 12th European Conference on Technology Enhanced Learning*, Tallin, Estonia, September, 2017, pp. 347–354.
26. D. Schultz, S. Duffield, S. C. Rasmussen and J. Wageman, Effects of the flipped classroom model on student performance for advanced placement high school chemistry students, *Journal of Chemical Education*, **91**(9), 2014, pp. 1334–1339.
27. B. Alden Rivers and J. T. E. Richardson, Illuminating the student experience within the liminal space: exploring data-driven learning design for negotiating troublesome concepts, *Higher Education Close Up 7 Conference*, Lancaster, UK, July 21–23, 2014.
28. L. Niles-Hofmann, *Data-Driven Learning Design. How to Decode Learner Digital Body Language*, 2016.
29. K. Mangaroska and M. Giannakos, Learning Analytics for Learning Design: Towards Evidence-Driven Decisions to Enhance Learning, *Proceedings of the 12th European Conference on Technology Enhanced Learning*, Tallin, Estonia, September, 2017, pp. 428–433.
30. A. Pardo, N. Mirriahi, R. Martinez-Maldonado, J. Jovanovic, S. Dawson and D. Gašević, Generating actionable predictive models of academic performance, *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, Vancouver, Canada, March 13–17, 2017, pp. 474–478.
31. J. Jovanović, D. Gašević, S. Dawson, A. Pardo and N. Mirriahi, Learning analytics to unveil learning strategies in a flipped classroom, *The Internet and Higher Education*, **33**, 2017, pp.74–85.
32. I. Estévez-Ayres, C. Alario-Hoyos, M. Pérez-Sanagustín, A. Pardo, R. M. Crespo-García, D. Leony, H. A. Parada and C. Delgado-Kloos, A methodology for improving active learning engineering courses with a large number of students and teachers through feedback gathering and iterative refinement, *International Journal of Technology and Design Education*, **25**(3), 2015, pp. 387–408.
33. B. A. Lindsey and M. L. Nagel, Do students know what they know? Exploring the accuracy of students' self-assessments, *Physical Review Special Topics-Physics Education Research*, **11**(2), 2015, pp. 020103:1–020103:11
34. E. Lahtinen, K. Ala-Mutka and H. M. Järvinen, A study of the difficulties of novice programmers, *ACM SIGCSE Bulletin*, **37**(3), 2005, pp. 14–18.
35. I. Milne and G. Rowe, Difficulties in learning and teaching programming—Views of Students and Tutors, *Education and Information technologies*, **7**(1), 2002, pp. 55–66.
36. E. Soloway and J. C. Spohrer, *Studying the Novice Programmer*, L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1988.
37. M. Grigoriadou, M. Toula and E. Kanidis, Design and evaluation of a cache memory simulation program, in Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies, Athens, Greece, July 9–11, 2003, pp. 170–174.
38. J. Sorva, *Visual program simulation in introductory programming education*. Ph.D. thesis, Aalto University, 2012.
39. A. Jimoyiannis and V. Komis, Computer simulations in physics teaching and learning: a case study on students' understanding of trajectory motion, *Computers & Education*, **36**(2), 2001, pp 183–204.
40. Virtualbox Project Homepage, https://www.virtualbox.org, Accessed 7 January 2017.
41. Apache Subversion Software Project Homepage, https://subversion.apache.org, Accessed 7 January 2017.
42. I. Estévez-Ayres, J. A. Fisteus and C. Delgado-Kloos, Lostrego: a distributed stream-based infrastructure for the real-time gathering and analysis of heterogeneous educational data, *Journal of Network and Computer Applications*, **100**, 2017, pp. 55–68.
43. J. A. Fisteus, N. Fernández García, L. Sánchez Fernández and D. Fuentes-Lorenzo, Ztreamy: A middleware for publishing semantic streams on the Web, *Web Semantics: Science, Services and Agents on the World Wide Web*, **25**, 2014, pp.16–23.
44. A. Martínez-Monés, Y. Dimitriadis, E. Gómez-Sánchez, B. Rubia-Avi, I. Jorrín-Abellán and J. A. Marcos, Studying participation networks in collaboration using mixed methods, *International Journal of Computer-Supported Collaborative Learning*, **1**(3), 2006, pp. 383–408.
45. W. He, A. Holton, G. Farkas and M. Warschauer, The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions, *Learning and Instruction*, **45**, 2016, pp. 61–71.
46. J. O'Flaherty and C. Phillips, The use of flipped classrooms in higher education: A scoping review, *The Internet and Higher Education*, **25**, 2015, pp. 85–95.
47. J. van der Meer, 'I don't really see where they're going with it': communicating purpose and rationale to first-year students, *Journal of Further and Higher Education*, **36**(1), 2012, pp 81–94.
48. A. Revell and E. Wainwright, What makes lectures 'unmissable'? Insights into teaching excellence and active learning, *Journal of Geography in Higher Education*, **33**(2), 2009, pp. 209–223.
49. B. Oakley, R. M. Felder, R. Brent and I. Elhajj, Turning student groups into effective teams, *Journal of Student Centered Learning*, **2**(1), 2004, pp. 9–34.
50. G. S. Mason, T. R. Shuman and K. E. Cook, Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course, *IEEE Transactions on Education*, **56**(4), 2013, pp. 430–435.
51. M. Dal, Teaching electric drives control course: Incorporation of active learning into the classroom, *IEEE Transactions on Education*, **56**(4), 2013, pp. 459–469.

Iria Estévez-Ayres is an Assistant Professor at the Telematic Engineering Department at Universidad Carlos III de Madrid. She obtained her Telecommunication Engineering degree from Universidad de Vigo in 2001 and her PhD from Universidad Carlos III de Madrid in 2007. She has been involved in several international and national projects and is author or co-author of more than 60 scientific publications in the fields of technology-enhanced learning, networking, distributed systems and real-time systems.

Jesus Arias Fisteus is Assistant Professor at the Telematic Engineering Department of Universidad Carlos III de Madrid (Spain). He holds a PhD degree in Communication Technologies from Universidad Carlos III de Madrid and a MSc degree in Telecommunication Engineering from the University of Vigo (Spain). Jesus has participated in several international and national research projects and is author of publications in top international journals and conferences in the fields of technology-enhanced learning, semantic web, networking and distributed systems.

Lucía Uguina-Gadella is a PhD Student at the Telematic Engineering Department of Universidad Carlos III de Madrid and also a Research Assistant in IMDEA Networks. She received her BchS degree in Telecommunication Technologies in 2016 and is ending her MS degree in Computational Engineering and Mathematics. Her research interests are Process Mining, Data Analytics and Learning Analytics, among others.

Carlos Alario-Hoyos is a Visiting Associate professor at the Telematic Engineering Department of Universidad Carlos III de Madrid. He received MS degree in Telecommunication Engineering and PhD in Information and Communication Technologies from Universidad of Valladolid, Spain, in 2007 and 2012 respectively. Carlos has received several awards for his work on educational technologies, including best paper award at EC-TEL 2013. He is author of more than 60 scientific publications and has participated in more than 20 research projects, in the regional, national and European levels. His skills and experience include research and development in MOOCs, social networks, collaborative learning, and evaluation of learning experiences, among others.

Carlos Delgado-Kloos received the PhD degree in Computer Science from the Technical University of Munich and in Telecommunications Engineering from the Technical University of Madrid. He is Full Professor at the Universidad Carlos III de Madrid, where he is Director of the UNESCO Chair on "Scalable Digital Education for All" and Vice President for Strategy and Digital Education.