Flipping Laboratory Sessions: An Experience in Computer Science

José A. Parejo, Javier Troya, Sergio Segura, *Member, IEEE*, Adela del-Río-Ortega, Antonio Gámez-Díaz, and Alfonso E. Márquez-Chamorro

***Abstract*— This paper reports our experience in flipping a second- year undergraduate course on software architecture and integration, taught in the second course of a Software Engineering degree. We compare the application of the flipped-classroom methodology with a traditional methodology. Our study encompasses two academic courses, in the years 2017 and 2018, and involves a total number of 434 students and 6 lecturers, placing this among the largest studies on flipped-classroom to date. The paper also reports on the production of the videos used with the flipped-classroom methodology, recorded by the lecturers in informal settings, and provides several lessons learned in this regard. The results of the study, backed by a solid statistical analysis of the data, demonstrate the suitability of the flipped-classroom methodology for laboratory sessions in the subject course. Among other results, our analysis concluded that students had on average 24 more minutes per session to solve in-class exercises with the flipped-classroom methodology; more than 70% of the students considered that the quantity, duration and didactic content of the videos were (very) appropriate; and 9 out of every 10 students would prefer this methodology in the laboratory sessions of future courses rather than a traditional face-to-face approach.**

***Index Terms*— Flipped classroom, software engineering, comparative study.**

I. Introduction

AN INCREASING number of papers [1], [8], [18] advocate the application of the flipped-classroom methodology conceived by Bergman and Sams in 2012 [4] as an option to optimally organize the learning time. This methodology consists in transferring to the student the responsibility of acquiring the most theoretical concepts before the class and devoting the in-person class to more practical activities such as solving exercises and discussions, where the assistance of the instructor is much more valuable. To this end, instructors typically transfer part of the content to videos that students must view before the face-to-face sessions.

This work was supported in part by the European Commis-sion (FEDER), in part by the Spanish Government under Project APOLO US-1264651, in part by the OPHELIA under Grant RTI2018–101204–B–C22, in part by HORATIO under Grant RTI2018-101204–B–C21, in part by the EKIPMENT-PLUS under Grant P18–FR–2895, and in part by the Program of Teaching Coordination and Innovation of the University of Seville 2017/2018 under Grant 21034. *(Corresponding author: José A. Parejo.)*

The authors are with the Department of Computing Languages and Systems, Universidad de Sevilla, 41012 Seville, Spain (e-mail: japarejo@us.es; jtroya@us.es; sergiosegura@us.es; adeladelrio@us.es; antoniogamez@us.es; amarquez6@us.es).

In the context of higher education, and in computer science in particular, many studies have evaluated the results of applying the flipped-classroom methodology both at the national level [12], [13], [17], [19], [20] as well as at the international level [3], [5], [14]–[16]. In most cases, there seems to be a positive effect of the application of the flipped-classroom on students' perception of the course and on their academic performance in it. Motivated by these results, the authors of this article proposed to solve a problem that was pressing in the course of Software Architecture and Integration, in the second year of the Software Engineering degree at the University of Seville. There was a lack of time to carry out exercises in laboratory sessions, partially due to the time invested to explain the technical details at the beginning of each lab. This explanation was aimed at technical concepts needed before starting to work on the lab exercises, such as the step-by-step configuration of the programming environment, instructions for cloud deployment, use of libraries, etc.

In this article, we present our experience after having applied the flipped-classroom methodology in the laboratory sessions of the above-mentioned course. We also describe in detail the data collection and analysis carried out and evaluate the impact on students. Our study differs from the existing ones in several aspects. Firstly, the number of students under study. In most cases this number is not very high, ranging from 12 [13] to 200 [3], [5], [15]–[17]. Only a small portion of the works so far has more than 300 students: 364 in [19] or almost 400 in [14], being our study one of the largest, with 434 students. Another aspect to take into account is the period of time to be considered, since in many cases only one course and year were considered, being it is difficult to compare the flipped-classroom with the traditional methodology. In contrast, our study encompasses two consecutive years in which we used the same contents and evaluation methods for a fair comparison of the results. It is also worth mentioning the large amount of data we collected, such as that related to class attendance, video views, perceived quality of the videos, number of exercises solved, duration of tutoring sessions, grades obtained, and degree of student satisfaction, among others. In total, we collected and analyzed about 2500 questionnaires and surveys, and the results of more than 1000 tests in the gamification platform Kahoot! [10], among other data, which allowed us to draw statistically significant conclusions. Another key aspect of our work, partly required by the volatility of the contents to be used in the course under study, was the need to seek a flexible approach to the methodology. Much of the content shown in the videos changes frequently, mainly due to the rapid evolution of the technologies used. This made us discard any process that could be slow and heavy in the recording of videos, so we propose a flexible guide for the elaboration of the videos, following existing recommendations [9].

Our study demonstrates the suitability of this methodology to be applied in the laboratory sessions of the Software Engineering area. Among other results, it is worth mentioning that: 1) 9 out of 10 students prefer the flipped-classroom methodology to the traditional one, 2) 86% of students watched the videos before attending the laboratory sessions, 3) we had an average of 24 minutes more per lab to devote to the execution of exercises as it was our goal, and 4) the duration of the tutoring sessions was reduced.

This paper is an extension of a previous paper [23] presented at the XXV "Jornadas sobre Enseñanza Universitaria de la Informática" (JENUI 2019), where it was selected among the six best papers of the conference to be sent to IEEE-RITA. The main extensions of this article with respect to the one presented in JENUI 2019 are the following:

• The subject and its context are described more extensively.

• The data collection is explained in more detail accompanied by 4 new figures. • A new research question has been introduced and answered, which investigates the influence of the selected methodology on the tutoring sessions.

The rest of the article is organized as follows. Section II describes the context of the course as well as the participants and methodologies followed in this study. Section III explains how data collection was carried out and Section IV shows how the video recording process was approached. The evaluation of the results is presented in Section V. Finally, the lessons learned and conclusions drawn from our work are presented in Sections VI and VII, respectively.

II. Context

The study was carried out in the course Software Architecture and Integration, covering two academic years, in 2017 and 2018, involving 6 instructors[1] and 434 students (225 in the first year and 209 in the second). Students were evaluated in the same way in the two years covered by the study. The traditional methodology was applied in the first year (2017), while the methodology based on flipped- classroom was applied in the second year (2018). Descriptions of the course, participants, methodology followed for the study, and data gathered are provided below.

A. The Course

The compulsory course Software Architecture and Integration is studied in the second semester (from February to June) of the second year of the Degree in Software Engineering at the University of Seville. The objectives of the course in general terms are (i) to understand the concepts of software architecture and the aspects that determine its design, (ii) to learn classic architectural patterns and styles, (iii) to create a simple architectural design document, (iv) to understand the concepts of application and data integration, (v) to identify and distinguish different integration patterns, and (vi) to develop hybrid web applications (so-called *mashups*) using web technologies such as REST and JavaScript. The course is taught over a period of about 15 weeks, with a theoretical and a practical part. The theoretical part consists of 18 face-to-face classes, while the practical part consists of 12 laboratory sessions in computer classes. Each class lasts 1 hour and 50 minutes.

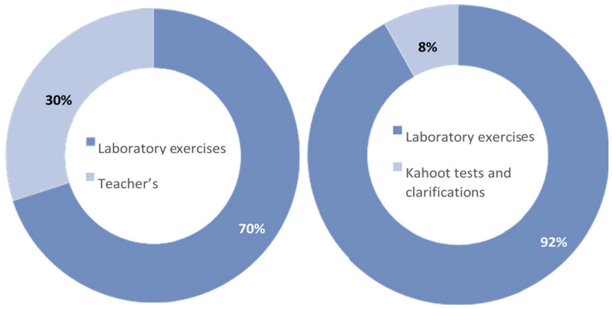
Students were evaluated in the same way in the two academic years. In the practical part, students had to organize themselves in groups of two to four students to carry out the course project, which consisted in developing a mashup (hybrid web application) by consuming third-party resources through their API REST (application interfaces). The mashup had to satisfy several requirements, including the integration of at least three real applications (through their APIs), the use of the model-view-controller architectural pattern, the modeling and documentation of the application architecture, and its deployment on the cloud. The project deliveries were organized in three incremental deliverables, evaluated by the instructors. Although the project was developed in groups, the final grade was individual, as usually not all students contribute equally to the project or have the same commitment and dedication. For this reason, the instructors had interviews with each group, where they asked each member of the group individual questions about the project developed. In addition, students had the possibility to take two optional programming exams during class hours, which could increase their final grade in the practical part up to 1 point, being 10 points the maximum achievable grade. As for the theoretical part, it was evaluated through a test at the end of the course. This had a weight of 40% in the final grade, while the grade obtained in the practical part (project plus optional programming exams) had a weight of 60%.

During both courses, instructors offered personalized tutoring sessions to groups to solve problems and doubts. These sessions had variable duration, usually depending on the difficulty of the problem to be solved and the knowledge of students about the problem. In order to have access to these tutoring sessions, students had to previously fill in a form in which they were asked for information about the problem they wanted to discuss, and they were asked for a diagnosis of its cause and a proposal of a possible solution. Therefore, before attending the tutoring session, students had to try to find out how their problems could be solved.

B. Participants

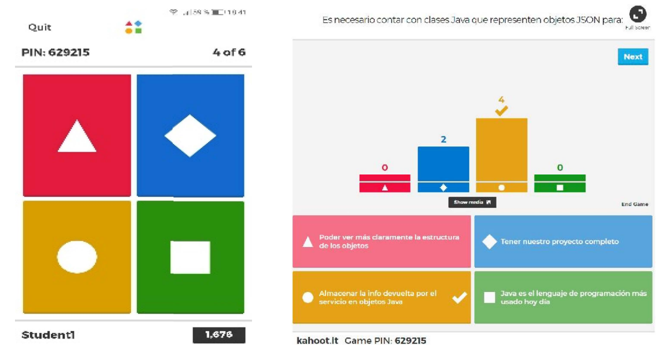
Six instructors participated in this study through the preparation and execution of the flipped-classroom methodology and data collection and analysis. The same group of instructors participated during the two academic years of the study, 2017 and 2018. The instructors had between 2 and 12 years of teaching experience in software engineering courses, and all had taught the course in previous years. As for students, the participants in the study were second-year undergraduates. A total of 434 students participated, 225 in the first year

[1]The 6 instructors are the authors of this paper. Throughout the paper, we write "the instructors" and "we" indistinctly



(a) Traditional methodology (b) Inverted class methodology

Fig. 1. Average time distribution.



(a) Answering on the mobile phone (b) Question result, displayed on the projector

Fig. 2. Using the Kahoot platform.

(2017) and 209 in the second (2018). Students were divided into 10 laboratory groups so that they could fit into the laboratories and receive appropriate attention from the instructors. Data collection from students was conveniently anonymized when requested, and none of the students requested to be excluded from the study.

C. Methodology

This study focuses on the comparison of the flipped-classroom methodology with traditional methodologies in the practical part of the subject. During the first year, 2017, a traditional methodology was followed, where at the beginning of each laboratory session the instructor explained everything related to practice, such as IDE (development environment) configuration, use of new libraries, steps to deploy applications on the cloud, etc. Slides, web pages, and code extracts were used for this purpose. These explanations ranged from 15 minutes (13.6% of the class time) to 55 (50% of the class time), with an average of 33.2 minutes (30% of the class time, Figure 1(a)). During the rest of the class, students had to do the proposed exercises and solve their doubts with the instructor. During the second year, 2018, we applied the flipped-classroom methodology. Thus, we recorded all the explanations on videos (see section IV), which students had to watch before attending class.

To find out if the concepts presented in the videos were understood by students, we used the gamification platform *Kahoot!*[1] to perform interactive tests with students at the beginning of the laboratory sessions. We displayed questions related to the video contents on the projector and students had to choose the correct answer using their mobile phone or computer (Figure 2(a)). Kahoot displays the number of correct answers after each question (Figure 2(b)) and ranks students according to the number of correct answers and the time it took them to answer. This encourages students to answer the questions as correctly and as fast as possible. Based on the answers obtained, the lab instructor decided which concepts to clarify. The Kahoot tests plus relevant clarifications took from 6 minutes (5.5% of class time) to 20 (18.2% of class time), with an average of 9.1 minutes (8.3% of class time, Figure 1(b)). This meant that students had an average of 24 more minutes than the previous year to do exercises and ask questions to the instructor.

[1]https://kahoot.com/

III. Data Collection

Throughout the two years, a large amount of data was collected for this comparative study. The data collected, its sources and measurement mechanisms are detailed below:

- *Grades obtained by students*. We got a total of 185 final grades in 2017 and 163 in 2018.
- *Duration of laboratory sessions.* As explained in the previous section, laboratory sessions worked differently in each academic course. For this reason, instructors wrote down the available time that students had to carry out exercises in all sessions.
- *Kahoot questionnaires*. The answers obtained in the Kahoot tests were stored after each class so that they could be part of the analysis. In order to have traceability of students, they introduced their digital identifier of the university as identifier in the Kahoot platform. In total, we stored 1027 answers to the Kahoot test questions.
- *Questionnaires on each laboratory*. To evaluate both methodologies, students completed a questionnaire after each lab session over the two years. They had to answer questions such as the number of exercises completed during the session, difficulty of exercises, interest in the lab session, clarity of explanations, technical quality of the videos (in 2018 only), etc. Students had to answer the questions using a Likert [2] scale, thus they were easy to answer, and this also facilitated further analysis of the answers. In total, we collected 1166 questionnaires in 2017 and 1174 in 2018.
- *Number and duration of tutoring sessions*. To evaluate the impact of the change of methodology in the tutoring sessions, instructors wrote down the duration of the session and the subject of students' doubts for all the sessions throughout the two years of the study. In addition, as mentioned above, to attend each tutoring session, students had to fill in a previous form. In this form, students were required to provide a minimal diagnosis of the problem and establish a possible cause for it. The data obtained from these answers were available to the instructors as soon as they were generated by students. Thus, doubts were answered more accurately and quickly during the tutoring session.
- *Final survey on the use of the flipped-classroom methodology:* Finally, at the end of the course in the

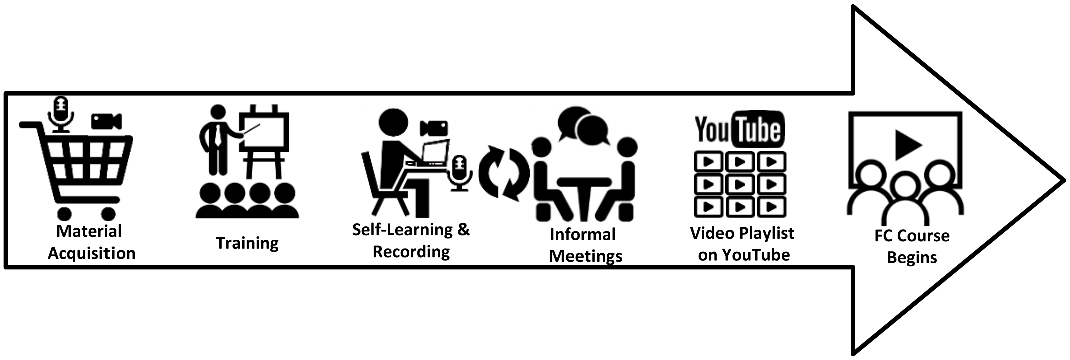


Fig. 3. Video recording and publishing process.

second year, students were asked to fill out a final survey showing their satisfaction with the flipped-classroom methodology. They completed it before knowing their final grade, so that this did not influence their answers. The survey included questions to be answered using a Likert scale, such as whether they watched the videos before class, whether they watched them at any other time during the semester, whether the length and content of the videos were adequate, etc. For other questions, it was allowed to write free text, for example about what parts of the videos were more useful, the preferred methodology for future classes, advantages and disadvantages of the flipped-classroom methodology, etc. In total, 144 surveys were collected.

All the information described above is available for consultation at [11]. In addition, the forms used for data collection and the scripts for statistical analysis of the information have been also included.

IV. Video Recording

To record and produce the videos, we first considered the services and resources offered by the university. The University of Seville offers a video recording service available to instructors[2]. To use this service, instructors must fill in a reservation form, wait for approval, go to a specific university building on the assigned date, prepare the recording contents and finally receive the edited video on a CD-ROM. This process was not adapted to the instructors' needs since it is a slow process and does not provide a flexible method to face the constant updates required by the agenda of the course. These changes, such as constant modifications of external API specifications and the tools used in the laboratory sessions, are unavoidable.

For this reason, it was essential for us to follow a flexible methodology for video recording, so that we could easily update the content of the videos at any moment. Therefore, we decided to record them ourselves, without any professional help. In fact, some recent studies have concluded that recording videos in informal settings can be even more effective than doing it in professional studios [9]. To this end, we needed to acquire the right material, as well as research about video recording and post-production, as only one of the instructors involved had previous experience. The following sections detail this process, which is summarized in Figure 3.

[2]https://servicio.us.es/websav/index.php/servicios/polimedia

A. Acquisition of Equipment

After researching the equipment needed for video recording, we purchased HD webcams, desktop and clip-on microphones, webcam tripods, mobile holders, and various cable adapters and extenders. The funding was obtained through a Teaching Innovation Project of the university and we were granted a budget of 2500€. The purpose of this project was the implementation of the flipped-classroom in the course, so the project was specifically assigned to us with this aim.

Unfortunately, the call for the innovation project expressly prohibited investing the money in the purchase of software licenses. This made impossible for us to obtain the latest version of popular video recording and editing tools, such as Camtasia or Adobe Premiere. This was a major obstacle to the recording and production of the videos. The solution was to use Camtasia Studio version 6 (released in 2008), available in our department, and the open OBS Studio (version 21).

B. Training and Preparation

The purpose of the videos was to capture the attention and interest of the student, so the way in which the videos were recorded was decisive. After studying guides and related work, we decided to follow the recommendations proposed by Guo *et al.* [9]. They presented an empirical study on how the way videos are produced affects students' commitments. In particular, we decided to establish a minimum set of guidelines to be followed for the creation of the videos:

1) *Videos have to be short.* According to Guo *et al.* [9], short videos are more interesting than long videos. As a result of the study, they recommend segmenting the agenda into videos of less than 6 minutes. Following this recommendation, we agreed to record up to 4 videos per lab session and no longer than 6 minutes.
2) *Instructors have to appear in the videos.* Guo *et al* [9] conclude that students prefer the face of instructors to appear in the videos since a human face gives a more 'intimate and personal' feeling and breaks the monotony of slides and code screenshots. They do not give any recommendations on how or when instructors should appear in the videos but do recommend that they appear whenever it makes sense in the video. We agreed to start all videos with the instructor by making an introduction to the video. Each instructor would decide when they

should appear in the rest of the video for any explanation or clarification.

3) *Videos must have good quality.* To ensure good quality videos, we organized a 4-hour course where basic concepts for flexible video recording and editing were explained. This course was taught by colleagues from our department with extensive experience in recording teaching videos. In the course, concepts such as the position of the instructor's body in the recording, the configuration of the microphone, video recording with mobile phones and tripods, or video editing, among others, were taught. We agreed that all videos should have a high-quality (Full HD resolution) and clear sound, avoiding any kind of echo or background noise.
4) *Spend time in pre-production.* Guo *et al.* [9] suggest that the pre-production phase (i.e. planning) has the greatest impact on the outcome of the videos, so they recommend investing effort in pre-production. Based on this, we all agreed to work hard in the preparation of each video, creating scripts for the videos, examples, and demonstrations, and trying to minimize improvisation.

Once we agreed upon these guidelines, each instructor became familiar with the purchased material and the recording and production software on their own. They researched and decided which configuration would be most appropriate. We held informal meetings to share our own findings so that everyone was aware of everything before and during video recording.

C. Recording and Result

Each instructor recorded the videos in her/his own office using clip-on microphones, webcams, and the necessary extension cables and adapters, as can be seen in Figure 4. In addition, two instructors also used a green chroma-key background to overlay content with post-production effects. Each instructor was responsible for recording the videos of one or two laboratory sessions. Some instructors could not fulfill the restriction of creating videos with a length fewer than 6 minutes. In total, we produced 29 videos, having the shortest a duration of 2 minutes 34 seconds and the longest 9 minutes 45 seconds. The average length was 5 minutes and 58 seconds, and we obtained 3.6 videos per lab session. All videos included the instructor introducing the lab session at the beginning of the video, and in some of them, the instructor also appeared in intermediate parts of the videos. An example of a video opening can be seen in Figure 5(a). One of the instructors decided to keep the upper body part throughout the video in the lower right corner, as shown in Figure 5(b).

All the videos were uploaded to a YouTube playlist[3]. The playlist was created as *unlisted*, so the videos did not appear in any Internet searches and only users with the link could access the videos. In this way, we avoided that people outside the course could access the videos, which could have compromised the results of the evaluation. We activated the option to allow comments on the videos, but they had to be

[3]http://bit.ly/PaqueteLaboratorioClaseInvertidaAISS

Fig. 4. Video recording in the office.

(a) Introduction to each video. (b) An instructor using a chroma-key background.

Fig. 5. Explanatory videos for the flipped-classroom.

reviewed and approved by the course instructors before being shown online.

V. Evaluation

In this paper, we aim to answer three research questions, namely:

- *RQ1 - What is the students' perception of the flipped-classroom methodology?*
- *RQ2 - What is the impact of the chosen methodology on student performance?*
- *RQ3 - What is the impact of the chosen methodology on the number and duration of tutoring sessions requested by students?*

Since we collected data from different sources, we use the correlation coefficient of *Pearson* to evaluate if there exists a linear relationship between two quantitative random variables. This coefficient shows how strong the relationship between the variables is. When this coefficient is not negligible, we also calculate the p-value, which determines whether the correlation is significant or not. Additionally, we compute the non-linear correlation coefficient of *Spearman* to determine monotonic correlations not identified by Pearson's coefficient. We also use the *Cronbach's Alpha* coefficient [7] to determine the reliability and consistency of the data provided by students in their final surveys on the flipped-classroom experience. Finally, regarding students' final grades and the duration of the tutoring sessions, since these data do not follow a normal distribution according to the results of the *Shapiro-Wilk* [22] test, we performed the test of signed ranges of *Wilcoxon* [24] to compare the mean range of two related samples and to determine whether there are differences between

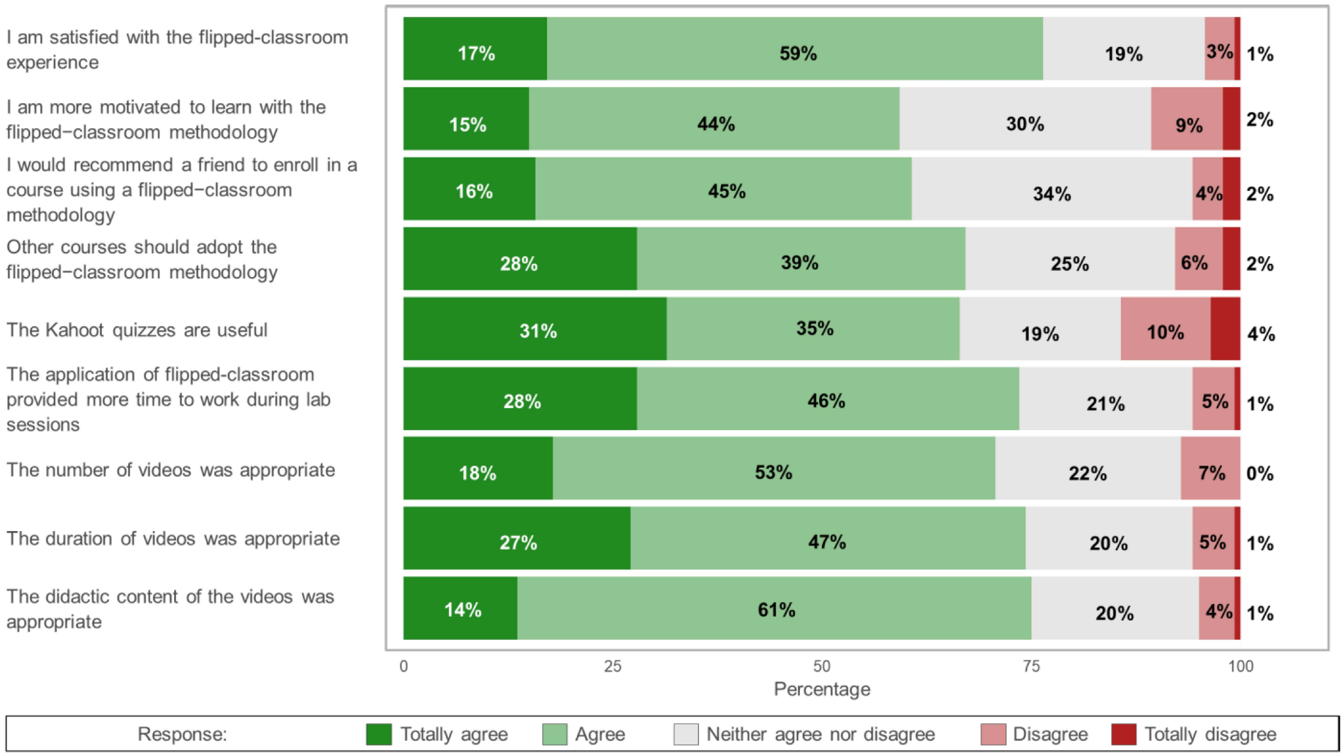


Fig. 6. Students' opinions about the flipped-classroom.

them. To evaluate the size of the effect associated with the differences between the measurements of the distributions, the Cohen's d estimator [6] was used.

A. RQ1 - Students' Perception

From the data analyzed, a decisive conclusion is that 87% of the students surveyed (122 out of 144) prefer the flipped-classroom methodology over traditional methodologies. Concerning the questions answered using a *Likert scale* [2] (*I totally agree, I agree, I neither agree nor disagree, I disagree, I totally disagree*), the answers related to the level of students' satisfaction are shown in Figure 6.

It should be noted that the responses were generally very favorable to the flipped-classroom methodology. Thus, 76% of the students surveyed were (very) satisfied with the flipped-classroom experience and 59% stated that they were motivated to learn more with this methodology. Besides, 61% and 67% of students would recommend a friend to enroll in a course that applies this methodology and thought that others should apply it, respectively. Three out of four students found the Kahoot tests useful.

More than half the students worked more at home with this methodology than with traditional methodologies. Concerning videos, 86% watched the videos before attending the practical sessions, and 64% watched them at other times during the semester. Figure 7 illustrates the percentage of students who claimed to have done some work prior to the laboratory sessions (reading the slides in the first year and watching the videos in the second). A very significant and homogeneous change is observed in every laboratory session. Finally, more than 70% of students were (very) satisfied with the number of videos, their duration, and their content. Most of the students

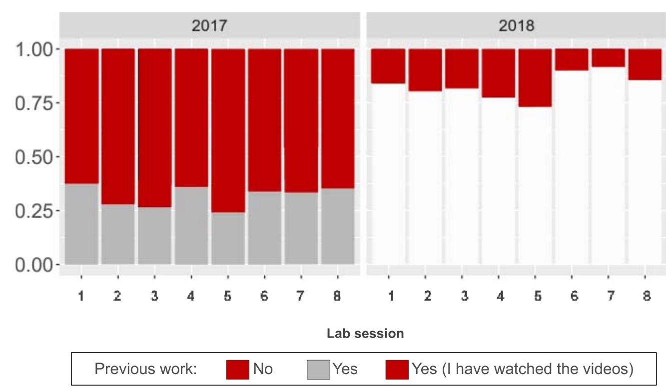


Fig. 7. Percentage of students who say they have done prior laboratory work by year and laboratory.

considered that the most valuable part of the videos are practical demonstrations, and, indeed, they requested these sections to be extended for future versions of the videos.

Figure 8 shows a violin plot describing the distribution of students' ratings regarding the technical quality of the videos in each laboratory session. These data were analyzed with the surveys collected after each class. The shape associated with the column for each laboratory session describes the probability density of occurrence of the different scale values (between 0 and 10) resulting from the students' rating. Furthermore, the plot represents the median and different quartiles with horizontal lines. As depicted in the figure, scores are quite good in general as the median values are around 7 in most cases. However, there are important variations between the different laboratory sessions, with laboratory 8 being the one that has obtained the worst results. Precisely, the videos of laboratory 8 hardly include practical demonstrations.

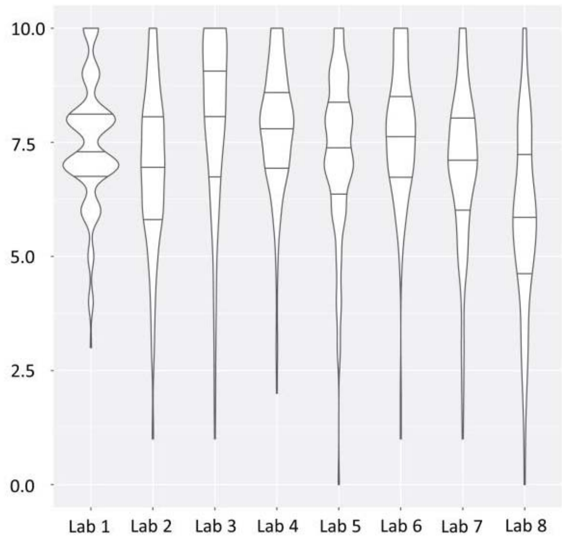


Fig. 8. Opinions on the technical quality of the videos.

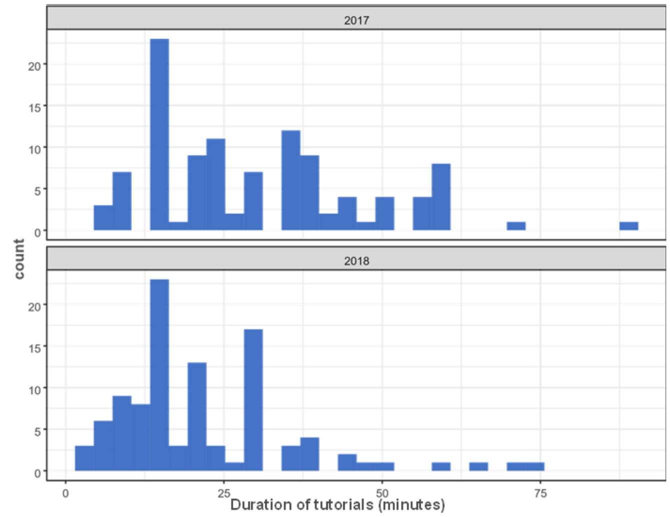


Fig. 9. Aligned histograms of the duration of tutoring sessions per year (in minutes).

B. RQ2 - Impact of the Methodology on Students' Performance

Results indicate that there is no evidence that the methodology applied influenced students' grades, either positively or negatively (the *Wilcoxon* test generated a *p-value* of 0.403).

Secondly, we studied if there exists any relationship between lab sessions attendance and obtained grades. To gather information on student's attendance, we simply needed to check whether they undertook the Kahoot test and/or submitted the questionnaire. In the first year, where the traditional methodology was applied, we obtained a positive correlation (with a linear correlation coefficient of 0.33), which means that students who attended classes usually obtained better grades. However, this correlation practically disappeared in the course where the flipped-classroom methodology was followed (with a linear correlation coefficient of 0.14, reducing it to less than half). This result is probably due to the fact that students can follow better the course when the flipped-classroom is applied, even if they cannot attend all laboratory sessions since they have the explanatory videos of every session available to watch at any time.

We also studied the relationship between attendance to previous laboratory sessions and the average number of exercises solved in each class. We obtained the same result: in 2017 there is a weak positive correlation (with a linear correlation coefficient of 0.23), while in 2018 it is greatly reduced until it almost disappears (the value of the correlation coefficient, in this case, was 0.09). Thus, in the first year, students who attended more classes managed to complete more exercises in subsequent laboratory sessions overall, while in 2018 attendance in previous classes was not as decisive in completing more exercises in subsequent laboratory sessions. In terms of the number of exercises completed in the laboratory sessions each year, in 2018, students finished 0.4 more exercises than in the previous year (representing up to 13.3% of the total exercises in some laboratory sessions). Moreover, in several cases, students did not do more exercises simply because they had finished them all. Thus, the percentage of students who finished all exercises increased from 22.4% in 2017 to 35.9% in 2018. This result is very positive, although predictable, as in 2018 they had on average 24 more minutes per class to do exercises (see Section II).

Finally, there is a slight positive correlation between students' final grades and the points obtained from the Kahoot tests (linear correlation coefficient of 0.28). This indicates that these tests were assessing how much students had learned after watching the videos, at least in the same sense as final grades of the course do. Furthermore, this learning was reinforced by taking Kahoot tests.

C. RQ3 - Impact of the Methodology on the Number and Duration of Tutoring Sessions

During the first year, 109 tutoring sessions were conducted, which means a total of 0.449 sessions per enrolled student. In the second year, the number of tutoring sessions decreased to 101, but considering the number of students enrolled that year, this means 0.483 tutoring sessions per student enrolled. These values are very similar, and we cannot affirm that the change in methodology had an impact on the number of tutoring sessions conducted. However, if we consider the duration of the tutoring sessions, significant differences do appear. The total time devoted to tutoring the first year was 3316 minutes, with an average duration of the sessions of 30.4 minutes. The total tutoring time the second year was 2217 minutes, with an average session length of 21.9 minutes. Figure 9 shows the aligned histograms of the duration of the tutoring sessions for both years. In the second year, there is a significantly higher number of shorter tutoring sessions (15 minutes or less) and the number of tutoring sessions lasting more than 35 minutes is significantly lower, even though there are still sporadic sessions of extremely long duration (up to 75 minutes).

The *Wilcoxon* test provides a *p-value* of 0.0000535, which allows us to say that the differences between the length of the sessions are significant. The average difference in session length is 7.5 minutes, which translates into a *Cohen's d-value* of 0.546. This, according to the interpretation guidelines for this estimator [21], translates into a *medium effect size*.

The reduction observed in the duration of the tutoring sessions may not be the exclusive consequence of the change in methodology, but rather the combination of the change in methodology and the use of a pre-session form. This form forces students to propose a hypothesis about the cause of the problem for which they want to attend the tutoring session and to make an initial search for the solution to the problem. The same form was used during the two years of this study. However, it is possible that during the solution-seeking phase, students used the videos developed for the flipped-classroom (and especially the practical demonstration sections). This might have helped them to pose the questions during the tutoring sessions in a clearer and more precise way, which resulted in a faster problem resolution and a shorter duration of the sessions. This hypothesis to explain the significant differences observed is supported by the fact that 64% of students claimed to have watched the videos at other times during the course (as explained in Section V-A), and by the students' assessment of the most useful sections of the videos (practical demonstrations, as mentioned in the following section).

VI. Lessons Learned

After our experience implementing the flipped-classroom methodology and the subsequent analysis we conducted, we are able to highlight several lessons learned regarding the video-recording process:

A. It Is Possible to Record Videos in a Flexible Way

While it is true that the learning curve was steep, the final results suggest that the effort invested is worthwhile. For example, it took time to find the right settings in the software to start recording the videos, as well as to handle all the post-production possibilities. However, once instructors learned the basics, we were able to produce high-quality 5-minute videos in about 20 minutes. We also learned that the minimum equipment for recording high-quality videos consists, apart from the software, of a webcam or mobile phone, a tripod, a lavalier microphone, and cable extenders. All this can be purchased from the Internet for less than 200 euros.

B. Post-Production Software Is Crucial

We realized that free video recording and production software often has some limitations. It was also a challenge to achieve the optimal configuration with the old version of Camtasia, as it is not designed for today's video resolutions and video formats. Therefore, we believe it is worthwhile to invest in video editing software.

C. Videos Have to Be Created Incrementally

We learned that it is better to record small video clips rather than long ones. Therefore, if mistakes are made on one of the clips during the recording, only that part needs to be redone. Similarly, if a mistake is identified during the post-production phase, only the small clip that contains the mistake needs to be redone. It is also useful to keep the videos year after year: several parts of the video can remain fixed, such as the opening and closing fragments, while intermediate parts will require to be updated to keep the examples shown in line with the updates performed on the software used at the course.

D. The Effort Involved in Applying the Flipped-Classroom Methodology Is Worthwhile

As we mentioned, producing the videos ourselves required a significant effort. Furthermore, preparing all Kahoot tests and the corresponding questionnaires also required several working days. However, the good organization between the instructors, initiated by the course coordinator, and the division of effort between all of them, made the whole process more manageable. Moreover, the videos and Kahoot tests can be reused in future editions of the course, which makes the effort worthwhile. As mentioned, some video parts would need to be remade for future use. For instance, in the following academic course (2019), all the videos and tests were reused as class material, although almost 40% of the videos had to be partially modified, either to improve the recording quality or to introduce modifications in the demonstrations due to the changes in the technology used.

E. Demos Are a Must

The lack of demos in some videos was criticized by students. We did not take this into account when designing the format for the videos, so we have set the goal of adding new demos to the videos of all laboratory sessions in subsequent editions of the course.

F. All Videos Must Have a Similar Format

All videos must follow the same format, as proposed by Guo *et al.* [9]. Although we tried to avoid it, our videos were somewhat heterogeneous. For example, two instructors used a chroma-key background on which they superimposed the content of the slides (see Figure 5(b)), while the other instructors did not use it, and the ratio of demos to slides is also unbalanced in the videos. The opinions expressed by students suggest that the degree of homogeneity of the videos is important and helps students to become familiar with them.

VII. Conclusion

This article presents a comparative study of the flipped-classroom methodology with a traditional methodology in a computer science course. The study covers two academic years and involves a total amount of 434 students. After the large amount of data collected and analyzed throughout the two academic years, the study concludes that 87% of students preferred the flipped-classroom methodology over traditional methodologies and that students had 24 more minutes to in-class work on the resolution of exercises. In addition, students seemed committed to this methodology, since 86% watched the videos before going to class, and more than 70% were satisfied with the number of videos, their length, and their content. The results also suggest that the application of the flipped-classroom methodology makes the course easier to follow by students who are not able to attend classes regularly. Finally, it was observed that tutoring sessions were shorter with the flipped-classroom methodology, probably because students could prepare prior to the sessions by watching video content.

This article also describes the methodology followed by the instructors for recording the videos used in the flipped-classroom methodology. In addition, several lessons learned are described, such as the importance of using demos in the videos and the quality of the recording software. This article lays the foundations for the application of the flipped-classroom methodology on a larger scale, including not only theory classes but also laboratory classes, even in courses with many students and highly volatile contents.

REPLICABILITY

For the sake of replicability of the study and verifiability of the results, a laboratory package has been created. It contains all the data collection forms used in this study, the data collected, and the scripts for statistical analysis of the data. These contents are available through the EXEMPLAR platform at [11].

REFERENCES

[1] L. Abeysekera and P. Dawson, "Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research," *Higher Edu. Res. Develop.*, vol. 34, no. 1, pp. 1–14, Jan. 2015.

[2] G. Albaum, "The Likert scale revisited," *Market Res. Society. J.*, vol. 39, no. 2, pp. 1–21, Mar. 1997.

[3] A. Amresh, R. A. Carberry, and Y. C. J. Femiani, "Evaluating the effectiveness of flipped classroom for teaching CS1," in *Proc. IEEE Frontiers Educ. Conf.*, vol. 27, Oct. 2013, pp. 733–735.

[4] B. F. Hantla, "Book review: Flip your classroom: Reach every student in every class every day.," *Christian Edu. J., Res. Educ. Ministry*, vol. 11, no. 1, pp. 183–188, May 2014.

[5] T. H.-C. Chiang, "Analysis of learning behavior in a flipped programing classroom adopting problem-solving strategies," *Interact. Learn. Environ.*, vol. 25, no. 2, pp. 189–202, Feb. 2017.

[6] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Evanston, IL, USA: Routledge, 2013.

[7] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, Sep. 1951.

[8] M. N. Giannakos and N. Chrisochoides, "Challenges and perspectives in an undergraduate flipped classroom experience: Looking through the lens of learning analytics," in *Proc. IEEE Frontiers Edu. Conf. (FIE)*, Oct. 2014, pp. 1–5.

[9] P. J. Guo, J. Kim, and R. Rubin, "How video production affects student engagement: An empirical study of MOOC videos," in *Proc. 1st ACM Conf. Learn. Scale Conf.*, 2014, pp. 41–50.

[10] (2003). *Kahoot! Online Platform*. [Online]. Available: https://kahoot.com

[11] (2019). *Experiment Materials*. [Online]. Available: https://exemplar.us.es/demo/Troya2019FlippedClassroom

[12] Y. M. Marcos. Á. L., "Una experiencia de clase invertida en la enseñanza de la programación," *Actas JENUI*, vol. 3, pp. 47–54, Mar. 2018.

[13] S. L. M. Y. E. S. Boró, "Mejora en el aprendizaje a través de la combinación de la clase invertida y la gamificación," *ActasJENUI*, vol. 2, pp. 213–220, Oct. 2017.

[14] M. L. Maher and C. Latulipe, "Flipped classroom strategies for CS education," in *Proc. 46th ACM Tech. Symp. Comp. Sci. Edu.*, vol. 15, 2015, pp. 218–223.

[15] 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 Trans. Educ.*, vol. 56, no. 4, pp. 430–435, Nov. 2013.

[16] K. K. Molnar, "What effect does flipping the classroom have on undergraduate student perceptions and grades?" *Edu. Inf. Technol.*, vol. 22, no. 6, pp. 2741–2765, Nov. 2017.

[17] J. Navarro, D. Amo, X. Canaleta, E. V. Vila, and Y. C. Martínez, "Utilizando analítica del apren- dizaje en una clase invertida: Experiencia de uso en la asignatura de Sistemas Digitales y Microprocesadores," *Actas JENUI*, vol. 3, pp. 391–394, Oct. 2018.

[18] J. O'Flaherty and C. Phillips, "The use of flipped classrooms in higher education: A scoping review," *Internet Higher Edu.*, vol. 25, pp. 85–95, Apr. 2015.

[19] A. P. Espinosa, B. P. Campos, and Y. B. D. P. Prieto, "Una experiencia de flipped class-room," *Actas JENUI*, vol. 12, pp. 237–244, Oct. 2016.

[20] S. R. Pascual, J. I. P. Navarrete, J. C. Soro, and M. P. Martínez, "De la clase tradicional a la clase invertida: Aplicación práctica en ingeniería del software," *En Actas de las JENUI*, vol. 3, pp. 119–126, Oct. 2018.

[21] S. S. Sawilowsky, "New effect size rules of thumb," *J. Mod. Appl. Stat. Methods*, vol. 8, no. 2, p. 26, 2009.

[22] S. S. Shakpiro and Y. Wilk, "An analysis of variance test for normality," *Biometrika*, vol. 52, nos. 3–4, pp. 591–611, Dec. 1965.

[23] J. Troya, S. Segura, A. José Parejo, A. D. Río-Ortega, A. Gámez-Díaz, and Y. E. A. Márquez-Chamorro, "Invirtiendo las clases de laboratorio en Ingeniería Informática: Un enfoque ágil," *Actas JENUI*, vol. 4, pp. 15–22, Oct. 2019.

[24] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.

José A. Parejo is currently a Professor with the University of Seville, Spain. He has seven years of experience in the software development industry and more than 12 years of experience teaching undergraduate and master courses on various topics related to software engineering. His research is focused on search-based software engineering and empirical software engineering.

Javier Troya received the Ph.D. degree in computer science from the University of Málaga, Spain, in 2013. He spent two and a half years working as a Post-Doctoral Researcher with TU Wien, Austria. He is currently an Associate Professor with the University of Seville, Spain, where he participates in several teaching innovation programs. His research focuses on the use of innovative methodologies for teaching, model-driven engineering, and software testing.

Sergio Segura (Member, IEEE) is currently a Professor with the University of Seville, Spain. He has more than 13 years of experience teaching undergraduate and master courses on various topics related to software engineering. He has supervised two doctoral theses, which were awarded. He seeks to improve student participation and motivation using innovative teaching techniques, such as inverted class and the use of serious games.

Adela del-Río-Ortega received the Ph.D. degree in software engineering from the University of Seville, Spain, in 2012. She is currently an Associate Professor with the University of Seville. Her research focuses on business process performance management. She did international research stays at HPI Potsdam, VU University Amsterdam, and UNIRIO. She has eight years of teaching experience in bachelor's and master's degrees in software engineering and information systems. She has supervised a thesis and is currently supervising a second one. She is also involved in several teaching innovation programs.

Antonio Gámez-Díaz received the B.S. degree in computer engineering in 2015 and the M.Sc. degree in software engineering and technology in 2016. He is currently a Pre-Doctoral Researcher with the University of Sevilla. He received the competitive Pre-Doctoral Fellowship (FPU) granted by the Spanish Government. He is a member of the ISA Research Group. His research interests include service-oriented computing and specifically in the government of restful systems. He collaborates with leading organizations such as the OpenAPI Initiative and Docker or GitHub Education, as an ambassador for the technological and industrial knowledge diffusion, having imparted several invited talks and workshops.

Alfonso E. Márquez-Chamorro is currently an Assistant Professor with the University of Seville, Spain, and is a member of the ISA Research Group. His current research areas include machine learning, business process management, and process mining. Previously, he worked in bioinformatics and evolutionary computing.