Informatics in Education, 2023, Vol. 22, No. 1, 121–139

DOI: 10.15388/infedu.2023.08

Integrating SPOCs in Software Testing Education: Evidence in Emergency Remote Courses

Jorge Marques PRATES[1], Silvana Morita MELO[2], Pedro Henrique Dias VALLE[3], Rogério Eduardo GARCIA[4], José Carlos MALDONADO[5]
[1]*State University of Mato Grosso do Sul – Dourados, Brazil*
[2]*Federal University of Grande Dourados – Dourados, Brazil*
[3]*Federal University of Juiz de Fora – Juiz de Fora, Brazil*
[4]*São Paulo State University – Presidente Prudente, Brazil*
[5]*University of São Paulo – São Carlos, Brazil*
e-mail: jprates@uems.br, silvanamelo@ufgd.edu.br, pedrovalle@ice.ufjf.br, rogerio.garcia@unesp.br, jcmaldon@icmc.usp.br

Received: December 2021

Abstract. Nowadays, SPOCs (Small Private Online Courses) have been used as complementary methods to support classroom teaching. SPOCs are courses that apply the usage of MOOCs (Massive Open Online Courses), combining classroom with online education, making them an exciting alternative for contexts such as emergency remote teaching. Although SPOCs have been continuously proposed in the software engineering teaching area, it is crucial to assess their practical applicability via measuring the effectiveness of this resource in the teaching-learning process. In this context, this paper aims to present an experimental evaluation to investigate the applicability of a SPOC in a Verification, Validation, and Software Testing course taught during the period of emergency remote education during the COVID-19 pandemic in Brazil. Therefore, we conducted a controlled experiment comparing alternative teaching through the application of a SPOC with teaching carried out via lectures. The comparison between the teaching methods is made by analyzing the students' performance during the solving of practical activities and essay questions on the content covered. In addition, we used questionnaires to analyze students' motivation during the course. Study results indicate an improvement in both motivation and performance of students participating in SPOC, which corroborates its applicability to the software testing teaching area.

Keywords: small private online courses, SPOCs, software testing, computing education.

1. Introduction

In 2020, the world was affected by the COVID-19 pandemic. According to UNESCO, over 100 countries have adopted sanitary measures, impacting a significant number of students UNESCO (2020). Measures against the spread of COVID-19 included the temporary closure of universities and the implementation of emergency remote teaching, changing the traditional teaching methods. In addition, during this period, online learning services had to be used to meet the needs of universities.

Discussions were held to find methods and means to adapt the current education system during the emergency period. Furthermore, this system's impact on learning was also evaluated by universities all over the world, mainly because most students were not familiar with the use of online platforms and their resources Wang *et al.* (2021). During the pandemic, the Brazilian authorities, respecting the biosafety guidelines and the determinations of the local health authorities, took isolation measures and suspended classroom teaching in all schools and universities. The alternative was to apply online teaching on a large scale through remote classes using video conferencing tools.

Among the alternatives, the SPOC (Small Private Online Courses) teaching and learning method has been adopted by colleges and universities worldwide Dong *et al.* (2021). SPOCs use the resources of MOOCs in private courses as supplementary material for classroom teaching Fox (2013). These courses are designed to serve a small number of students Muñoz-Merino *et al.* (2017), which makes the interactions among the participants and instructors easier.

This study presents an experimental study investigating the application of a SPOC in the teaching of software testing during emergency remote teaching due to the COVID-19 pandemic. We conducted a controlled experiment to compare alternative teaching through the application of the SPOC with teaching carried out through online lectures. The comparison was performed on teaching methods, motivation, and student performance, while students accomplished Verification, Validation, and Software Testing activities. The results obtained indicate better motivation and performance results of students participating in the SPOC.

This article is organized as follows. The following section provides an overview of teaching computing and software testing with SPOCs. In section 3, the research method through a controlled experiment is described. The results of the controlled experiment are presented and discussed in section 4. We conclude this article in the last section.

2. Small Private Online Courses

A MOOC (Massive Open Online Course) is a web-based distance course available to many geographically distributed students. From the MOOCs, variations of courses offered to students emerged, such as the SPOCs, made available in most cases by universities to internal and external audiences. Unlike MOOCs, SPOCs are private courses offered to few people. Their use integrates online teaching with the traditional small-scale classroom to support classroom or applied classes in conjunction with methodologies such as flipped classroom Mutawa (2016); Wang (2017). In addition, students can access video lessons and other materials before they meet to discuss and solve problems under the guidance of teachers or instructors Fox (2013); Kaplan and Haenlein (2016).

SPOC resources include online lectures, readings, forums, and assessment activities Cheng and Zhang (2014). Online lectures can be delivered via video lessons, lasting between 5 and 10 minutes. During the week, students watch approximately one hour of video lessons, which they can watch at their own pace. In addition, there are two types of readings, the materials that are part of the course and the additional readings, both available on the online learning platform An *et al.* (2017).

In the forums, discussions related to the learning contents of each week are provided, which are evaluated by activities. Assessment activities are usually applied as weekly quizzes but can also be online and offline, such as hands-on assignments. Students can ask questions during the learning process by sending messages to their teachers and instructors, who must be prepared to answer the questions and interact with students Zhou *et al.* (2016); Filius *et al.* (2018). SPOCs are also applied with teaching methodologies such as the flipped classroom or blended learning Martínez-Muñoz and Pulido (2015); Jong (2016); Chen and Zheng (2017); Alario-Hoyos *et al.* (2017).

SPOCs have been applied in computing teaching, especially in studies about programming language and computer principles teaching Piccioni *et al.* (2014); Chu *et al.* (2015); Prates *et al.* (2019). The work of Li and Gu (2020) adopted a blended teaching method applying SPOCs in programming teaching. Classroom and online teaching were combined: students accessed the course content on the course website before the classroom. A mixed teaching method was designed by Zhou *et al.* (2019), which combined classroom with online teaching, claiming to impact students' performance and programming ability. In the study presented by Lui *et al.* (2017), the use of a SPOC and animated videos were used for project management teaching. The results show that learning motivation positively influenced students' performance in learning through a SPOC.

Teaching software testing with SPOCs is applied using blended learning and flipped classrooms. For example, the work of Liu (2020) combined online and traditional teaching using a blending teaching mode for software testing courses in three aspects: teaching method, teaching content, and teaching evaluation. Students used platform resources such as online lessons, tests, homework, and other offline resources to complete classes while communicating and discussing their content.

3. Experimental Study

An experiment is an experimental investigation that manipulates a factor or variable in the study environment. Experiments are conducted when it is essential to control a situation and manage behavior directly, precisely, and systematically. Besides, experiments involve more than one treatment to compare results, and they can allow the generalization of these results to a larger population Wohlin *et al.* (2012). In this study, an experiment was conducted to check the application of SPOCs in Software Testing Education. Additionally, it provides means for replicating the experiment in other domains of software engineering teaching.

This experimental study aims to evaluate the efficiency of applying SPOCs during the period of an emergency remote teaching in Software Testing education. For this, the alternative education through the application of SPOC is compared with traditional education (online lectures), applied to students in Software Testing education. The comparison is carried out by analyzing their performance and motivation.

The definition of which data is collected was based on the method GQM (*Goal/Question/Metric*) Basili *et al.* (1994), as presented as follows:

- **Study entity:** The object of study is the application of a SPOC.
- **Focus:** Efficiency and motivation.
- **Context:** The study was applied with a group of students from the Computer Engineering course at the Federal University of Grande Dourados. The participants are students who were freshmen until senior.
- **Specific objectives:**

  Analyze application of a SPOC
  for the purpose of evaluate it
  with respect to performance and motivation
  from the point of view of the undergraduate students
  in the context of undergraduate courses in the computing field.

To conduct the experiment, research questions and the metrics used were defined as shown below.

- **Research Questions:**
  - **RQ1:** Does the application of SPOC provide greater motivation during student learning compared to traditional education?
  - **RQ2:** Is there a difference in students' performance who participated in education using SPOC concerning students who participate in traditional education?
- **Metrics:**
  - **Performance:** the activities and multiple-choice tests applied to both courses (SPOC and traditional) are analyzed.
  - **Motivation:** the answers to the CIS and the IMMS questionnaires are analyzed to assess students' motivation that received the classes with SPOC and who received the traditional classes.

The CIS (*Course Interest Survey*) and IMMS (*Instructional Materials Motivation Survey*) questionnaire are measurement tools integrated in the ARCS model proposed by Keller Keller (2010). According to the ARCS model, four components affect motivation in the learning process: attention, relevance, confidence, and satisfaction. The CIS and IMMS questionnaires were designed to support the measurement of these motivational components.

The CIS questionnaire is designed to measure student reactions in a specific learning environment facilitated by instructors. There are 34 statements in this questionnaire distributed in four concepts: attention, relevance, confidence, and satisfaction. The survey can be composed of the average score on each of the four concepts or calculating

Table 1
Guide of CIS and IMMS Score

Dimension	CIS – Questions	IMMS – Questions
Attention	1, 4, 10, 15, 21, 24, 26 (R), 29	2, 8, 11, 12 (R), 15 (R), 17, 20, 22 (R), 24, 28, 29 (R), 31 (R)
Relevance	2, 5, 8 (R), 13, 20, 22, 23, 25 (R), 28	6, 9, 10, 16, 18, 23, 26 (R), 30, 33
Confidence	3, 6 (R), 9, 11 (R), 17 (R), 27, 30, 34	1, 3 (R), 4, 7 (R), 13, 19, 25, 34 (R), 35
Satisfaction	7 (R), 12, 14, 16, 18, 19, 31 (R), 32, 33	5, 14, 21, 27, 32, 36

the total score. The response scale ranges from 1 to 5, the minimum survey score is 34, and the maximum is 170.

However, each concept's minimum, medium, and maximum points vary because not all have the same number of items. Besides, some questions have a negative scoring scale, as shown in Table 1. Therefore, response values must be reversed before adding to the response total. That is, for these items, 5 –> 1, 4 –> 2, 3 –> 3, 2 –> 4 and 1 –> 5.

The IMMS questionnaire was designed to measure students' motivation regarding instructional materials. There are 36 statements in this questionnaire distributed in 4 concepts: attention, relevance, confidence, and satisfaction. The survey can be scored by finding the average score on each of the four concepts or calculating the total score. The response scale ranges from 1 to 5, the minimum survey score is 36, and the maximum is 180. The minimum, medium, and maximum points for each concept can vary in the CIS questionnaire. Besides, some questions have a negative scoring scale. Thus, answers must be reversed before being added to the total of answers: 5 = 1, 4 = 2, 3 = 3, 2 = 4 and 1 = 5.

3.1. *Planning*

Hypotheses are essential tools for experimentation. A hypothesis is formally defined, and data collected during the experiment can reject the hypothesis when possible. Conclusions can be made based on the hypothesis testing whether the hypothesis was rejected Wohlin *et al.* (2012). Three hypotheses were formulated: (i) null hypothesis – H_0; (ii) alternative hypotheses – H_1 and H_2. A null hypothesis represents no trends or differences among the phenomena measured in the experiment environment. On the other hand, an alternative hypothesis represents the hypothesis in which the null hypothesis is rejected. In this context, the hypotheses established for the experiment were:

- **RQ$_1$:** Does the application of SPOC provide greater motivation during student learning when compared to traditional education?
  - Null Hypothesis (H_0): Motivation (SPOC) = motivation (traditional education). In other words, there was no significant difference in motivation among students participating in SPOCs and in the traditional classes.

- ○ Alternative Hypothesis (H_1): Motivation (SPOC) > Motivation (traditional education). In other words, the student's motivation to participate in SPOCs is significantly greater than that of students participating in traditional classes.
- ○ Alternative Hypothesis (H_2): Motivation (SPOC) < Motivation (traditional education), in other words, the student's motivation to participate in SPOCs is significantly worse than that of students participating in traditional classes.

- **RQ2:** Is there a difference in students' performance who participated in education using SPOC concerning students who participate in traditional education?
  - ○ Null Hypothesis (H_0): Performance (SPOC) = performance (traditional education). In other words, there was no significant difference in performance among students in the classes with SPOC and in the traditional classes.
  - ○ Alternative Hypothesis (H_1): Performance (SPOC) > Performance (traditional education), in other words, the student's performance in classes with SPOC is significantly greater than that of students participating in traditional classes.
  - ○ Alternative Hypothesis (H_2): Performance (SPOC) < performance (traditional education), in other words, the student's performance in classes with SPOC is significantly worse than that of students participating in traditional classes.

Independent and dependent variables were selected to evaluate which of the defined hypotheses were confirmed. Independent variables can be controlled and changed in the experiment, while the dependent variables represent the effects of the treatments applied during the experiment. Fig. 1 shows the dependent and independent variables defined for this experiment.

The software testing education is the independent variable investigated, and it must be manipulated during the experiment. The dependent variables are related to the results obtained in the experiment. In this context, students' motivation is measured through the scores obtained in the CIS and IMMS questionnaires, while students' performance is evaluated through tests carried out at the end of each course module. Careful selection of experiment participants is essential because it is essential for generalizing the results. Therefore, undergraduate students from the Computer Engineering undergraduate course at the Federal University of Grande Dourados were invited to perform this study.

It is worth highlighting that all students participated in the experiment voluntarily and confirmed their interest in participating in the study by signing a consent form. In addition, the students also fill out a profile characterization form to verify their knowledge level in Software Testing and the Python language.

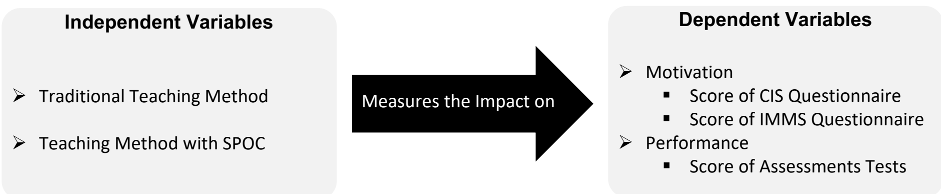

Fig. 1. Selection of experiment variables.

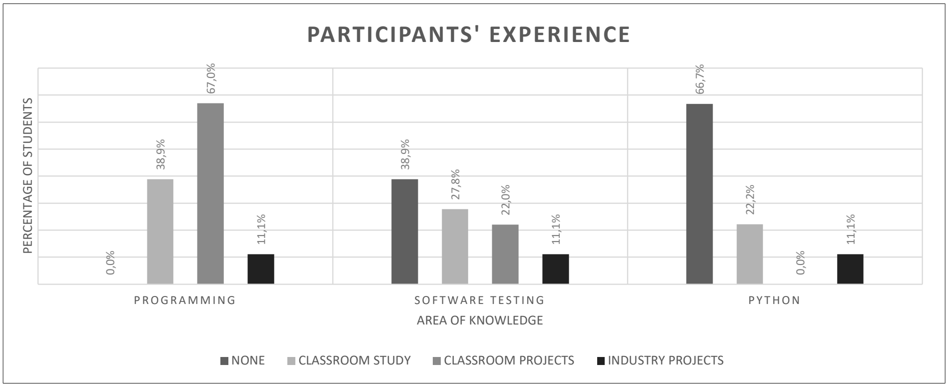

Fig. 2. Participants' level of experience.

Fig. 2 summarizes the participants' level of experience regarding the areas of knowledge that involve the study. Few students have applied programming concepts in real projects, while most know the programming from studies and projects developed in the classes. However, all students know some programming language.

Regarding participants' experience in software testing, most of them do not know software testing, and few have applied their knowledge in industrial projects. Finally, about the participants' experiences in the Python language, most of them have no language experience. A small part of the participants applied their knowledge to industrial projects. The other ones presented an experience in Python language through studies performed in the classes.

3.2. *Design*

The experiment design describes how tests are organized and performed. This experiment is defined with a factor and two treatments. In particular, the education with SPOC was compared with traditional education regarding the software testing education fundamentals. Participants were distributed balanced, and the division of groups was defined based on the analysis of responses obtained in the profile characterization form. The distribution of participants was organized as shown in Table 2.

Fig. 3 shows the plan followed to experiment. In the first step, the profile characterization form was applied to divide the participants into the control and experimental

Table 2

Distribution of participants

Groups	Traditional Education	Education with SPOC
Group 1	X	
Group 2		X

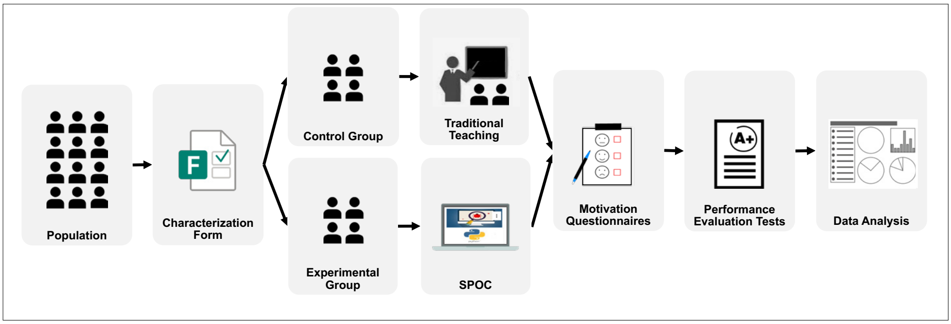

Fig. 3. Process to perform the study.

groups. Then, the groups of students were divided. The experimental group participated in the education with SPOC using the content addressed in the course "Introduction to Software Testing in Python Language". On the other hand, the control group had access to the same content through traditional classes with the support of material posted on the *Moodle* platform, simulating an education environment in the class but virtually.

After conduced the previous steps, the participants answered the CIS and IMMS questionnaires, evaluating the course's interest criteria. Additionally, assessments at the end of each module were made available to both groups with questions and practical activities that address the content considered in the two courses.

4. Results and Discussion

The collected data from the experimental study are used in the analysis and interpretation stage to conclude the investigation. The conclusions of each study research question are presented below.

Q1. Did SPOC application provide greater motivation during student learning when compared to traditional teaching?

The first research question aims to investigate the participant's motivation. The results of the CIS questionnaire application which assess motivation related to the course can be seen in Tables 3, 4, 5 and 6.

Attention was the first dimension evaluated, in which the experimental group achieved better results in all aspects compared to the control group. The results suggest that the participants of the SPOC were moderately attentive to the course and more motivated by the topics and problems presented in the course.

The experimental group achieved equal or superior results in all evaluated items in the relevance dimension. The results indicate that SPOC participants consider the course to be relevant. However, the interaction between the participants through the forums and tools did not reach a good evaluation.

Table 3
The CIS Questionnaire: Attention Dimension

Dimension: Attention	Mean: Experimental Group	Mean: Control Group
1. The instructor knows how to make us feel enthusiastic about the subject matter of this course.	2.83	2.33
4. This class has very little in it that captures my attention. (reverse).	2.33	3.50
10. The instructor makes the subject matter of this course seem important.	2.50	2.33
15. The students in this class seem curious about the subject matter.	2.83	2.17
21. The instructor does unusual or surprising things that are interesting.	2.50	1.67
24. The instructor uses an interesting variety of teaching techniques.	2.83	1.67
26. I often daydream while in this class. (reverse)	2.00	3.83
29. My curiosity is often stimulated by the questions asked or the problems given on the subject matter in this class.	3.00	1.83

Table 4
The CIS Questionnaire: Relevance Dimension

Dimension: Relevance	Mean: Experimental Group	Mean: Control Group
2. The personal benefits of this course are clear to me.	4.17	4.17
5. The instructor makes the subject matter of this course seem important.	3.17	2.67
8. I do NOT see how the content of this course relates to anything I already know. (reverse)	1.33	1.33
13. In this class, I try to set and achieve high standards of excellence.	3.83	2.67
20. The content of this course relates to my expectations and goals.	3.33	3.33
22. The students actively participate in this class through forums and other means of interaction.	1.83	1.83
23. To accomplish my goals, it is important that I do well in this course.	4.00	3.67
25. I do NOT think I will benefit much from this course. (reverse)	1.83	1.83
28. The personal benefits of this course are clear to me.	3.33	2.83

The experimental group also achieved equal or superior results in all evaluated items regarding the satisfaction dimension. The results indicate that SPOC participants expressed satisfaction in taking the course, mainly supported by questions 16 and 31. However, the low evaluation of the criterion related to comments and *feedback* indicates that it is something that must be improved in the course. In addition, it can be noted that it is a factor that is related to the low adherence to the discussion forums.

Confidence is the last dimension evaluated, in which the experimental group achieved superior results in most of the evaluated items. The results indicate that the SPOC participants were confident when taking the course, especially if they expended enough effort.

Table 5

The CIS Questionnaire: Satisfaction Dimension

Dimension: Satisfaction	Mean: Experimental Group	Mean: Control Group
7. I have to work too hard to succeed in this course. (reverse)	2.50	3.17
12. I feel that this course gives me a lot of satisfaction.	3.50	2.83
14. I feel that the grades or other recognition I receive are fair compared to other students.	3.67	3.00
16. I enjoy working on this course.	3.67	3.67
18. I get enough feedback to know how well I am doing.	3.33	2.83
19. I feel that this course gives me a lot of satisfaction.	3.50	2.83
31. I feel rather disappointed with this course. (reverse)	2.33	2.67
32. I feel that I get enough recognition for my work in this course by means of grades, comments, or other feedback.	2.67	1.50
33. The amount of work I have to do is appropriate for this type of course.	3.50	3.67

Table 6

The CIS Questionnaire: Confidence Dimension

Dimension: Confidence	Mean: Experimental Group	Mean: Control Group
3. I feel confident that I will do well in this course.	3.50	3.33
6. You have to be lucky to get good grades in this course. (reverse)	2.17	2.66
9. Whether or not I succeed in this course is up to me.	3.83	4.00
11. The subject matter of this course is just too difficult for me. (reverse)	3.16	2.33
17. It is difficult to predict what grade the instructor will give my assignments. (reverse)	3.83	3.50
27. As I am taking this class, I believe that I can succeed if I try hard enough.	4.17	3.67
30. I find the challenge level in this course to be about right: neither too easy not too hard.	3.17	3.50
34. I get enough feedback to know how well I am doing.	1.83	1.33

Table 7

The CIS Questionnaire: score of 4 dimensions

Evaluation Item	Mean: Experimental Group	Mean: Control Group
Attention	2.98	2.02
Relevance	3.59	3.33
Confidence	3.21	3.04
Satisfaction	3.18	2.91

However, the results suggest that the criterion related to *feedback* also needs to be improved in the course.

The comparison between the four dimensions of the two groups is shown in Table 7. One may note that, in general, the SPOC participants showed greater motivation at all levels, especially in the attention dimension.

Participants' motivation was also assessed using the IMMS questionnaire. The results, which assess motivation in relation to the course materials, are presented in Tables 8, 9, 10 and 11.

The experimental group achieved better results in almost all aspects of the attention dimension than the control group. The results suggest that the SPOC participants were attentive to the material used in the course. The approach adopted and the quality of the writing of the didactic and instructional materials used in the SPOC.

In the relevance dimension, the experimental group achieved superior results in most evaluated items. The results indicate that SPOC participants consider the course material and content relevant. Furthermore, it is possible to observe that the feeling of importance upon completing the course is highlighted in the evaluation.

The experimental group also achieved superior results in all the evaluated items in the satisfaction dimension. The results indicate that SPOC participants were satisfied with the course content, especially when completing the proposed activities and the course as a whole.

Finally, the experimental group achieved equal or superior results in the evaluated items in the confidence dimension. The results suggest that SPOC demonstrated confidence in the course content and material. Participants were confident that they could learn the content covered, as shown in the results of question 13.

Table 8
The IMMS Questionnaire: Attention Dimension

Dimension: Attention		Mean: Experimental Group	Mean: Control Group
2.	There was something interesting at the beginning of this course that got my attention.	3.83	2.83
8.	These materials are eye-catching.	3.67	2.5
11.	The quality of the writing helped to hold my attention.	4.00	2.00
12.	This course is so abstract that it was hard to keep my attention. (Reverse)	1.67	3.83
15.	The style of course is boring. (Reverse)	1.83	4.00
17.	The way the course information is arranged helped keep my attention.	3.67	2.33
20.	This course has things that stimulated my curiosity.	3.33	3.00
22.	The amount of repetition in this course caused me to get bored sometimes. (Reverse)	2.00	3.33
24.	I learned some things that were surprising or unexpected.	3.67	3.67
28.	The variety of reading passages, exercises, illustrations, etc., helped keep my attention on the course.	3.67	2.17
29.	The style of teaching materials writing is boring. (Reverse)	1.83	3.50
31.	There are so many words on each slide that it is irritating. (Reverse)	1.67	2.17

Table 9

The IMMS Questionnaire: Relevance Dimension

Dimension: Relevance	Mean: Experimental Group	Mean: Control Group
6. It is clear to me how the content of this material is related to things I already know.	3.50	3.33
9. There were a lot of examples that showed me how this material could be important for people that are learning about Software Testing.	3.50	2.67
10. Completing this course successfully was important to me.	4.17	3.67
16. The content of this material is relevant to my interests.	3.50	4.00
18. There are explanations or examples of how people use the knowledge in this course.	3.5	2.17
23. The content and style of writing in this course convey the impression that its content is worth knowing.	3.83	2.83
26. This course was not relevant to my needs because I already knew most of its content. (Reverse)	1.50	2.17
30. I could relate the content of this course to things I have seen, done, or thought about in my professional career.	3.33	2.50
33. The content of this course will be useful to me.	4.33	3.50

Table 10

The IMMS Questionnaire: Satisfaction Dimension

Dimension: Satisfaction	Mean: Experimental Group	Mean: Control Group
5. Completing the exercises in this course gave me a satisfying feeling of accomplishment.	3.67	3.00
14. I enjoyed this course so much that I would like to know more about this topic.	3.33	2.33
21. I really enjoyed studying this course.	3.50	2.67
27. The feedback after the exercises or other comments in this course helped me feel rewarded for my effort.	3.50	2.33
32. I felt good about completing this course successfully.	4.00	3.33
36. It was a pleasure to work on such a well-designed course.	3.83	3.00

The comparison of the IMMS questionnaire's four dimensions between the two groups is presented in Table 12. It is possible to observe that the SPOC participants had greater motivation in all dimensions, mainly in the attention dimension.

The purpose of the hypothesis test is to verify if it is possible to reject a particular null hypothesis based on a sample of some statistical distribution Wohlin *et al.* (2012). The *t-test* is a parametric test used to compare two independent samples, i.e., the design of the experiment must be one-factor with two treatments.

To check samples normality and, consequently, the *t-test* applicability, Shapiro–Wilks test was performed. For the control group sample, *p-value* = 0.666514 was ob-

Table 11

The IMMS Questionnaire: Confidence Dimension

Dimension: Confidence	Mean: Experimental Group	Mean: Control Group
1. When I first looked at this course, I had the impression that it would be easy for me.	3.33	3.33
3. This course was more difficult to understand than I would like for it to be. (Reverse)	2.5	3.17
4. After reading the introductory information, I felt confident that I knew what I was supposed to learn from this course.	3.17	2.67
7. Many of the pages had so much information that it was hard to pick out and remember the important points. (Reverse)	2.00	3.50
13. As I worked on this course, I was confident that I could learn the content.	3.83	2.67
19. The exercises in this course were too difficult. (Reverse)	2.83	3.50
25. After working on this course for a while, I was confident that I would be able to pass a test about Software Testing.	3.50	2.17
34. I could not really understand quite a bit of the material in this course. (Reverse)	3.33	3.67
35. The good organization of the content helped me be confident that I would learn this course.	3.50	3.00

Table 12

The IMMS Questionnaire: score of 4 dimensions

Evaluation Item	Mean: Experimental Group	Mean: Control Group
Attention	3.90	2.64
Relevance	3.79	3.17
Confidence	3.64	2.78
Satisfaction	3.41	2.67

tained, while for the experimental group sample, *p-value* = 0.899189. In both cases the *p-value* > 0.05 is obtained, so it is possible to consider that the samples are normal.

Assuming the significance value of $\alpha = 0.01$, the values obtained by applying *t-test* are: *t-value* = 3.52433 and $p = 0.003368$. Thus, is possible to reject the null hypothesis with a two-tailed test at $p < 0.01$.

Q2. Were there differences in the students' performance who participated in the SPOC concerning students who participated in a traditional class?

Fig. 4 shows the students' performance belonging to control and experimental groups in each of the course modules. It is noted that the students performance of SPOC participants, measured through tests and performance in hands-on activities, was better in all topics of the course.

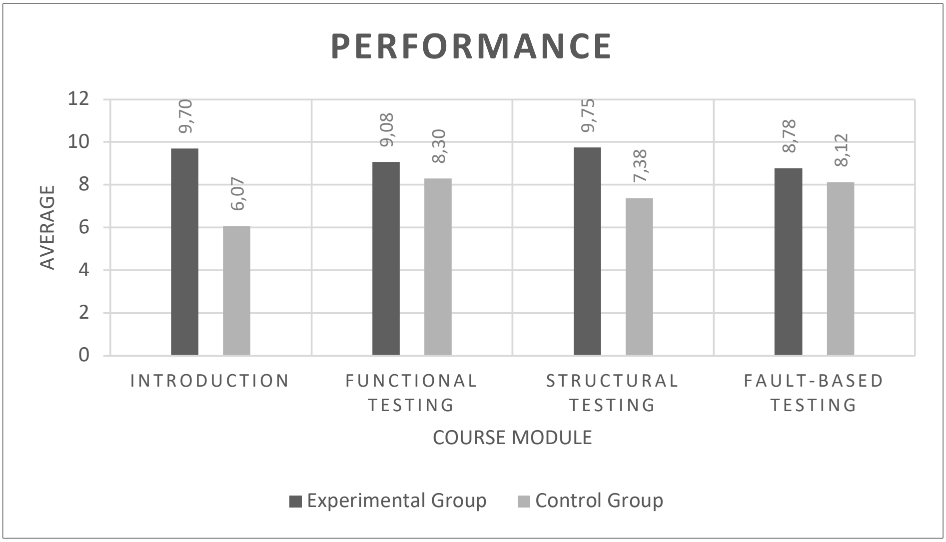

Fig. 4. Performance of control and experimental groups.

Fig. 5 shows a *boxplot* (box diagram), used in exploratory analysis of quantitative variables. Through the diagram, which represents the control group's performance in the experimental study, it is possible to observe that the data dispersion (difference between the first and quartiles) is small in the first three topics of the course. However, on the topic of "Defect-Based Testing," there was a greater discrepancy in the participants' performance.

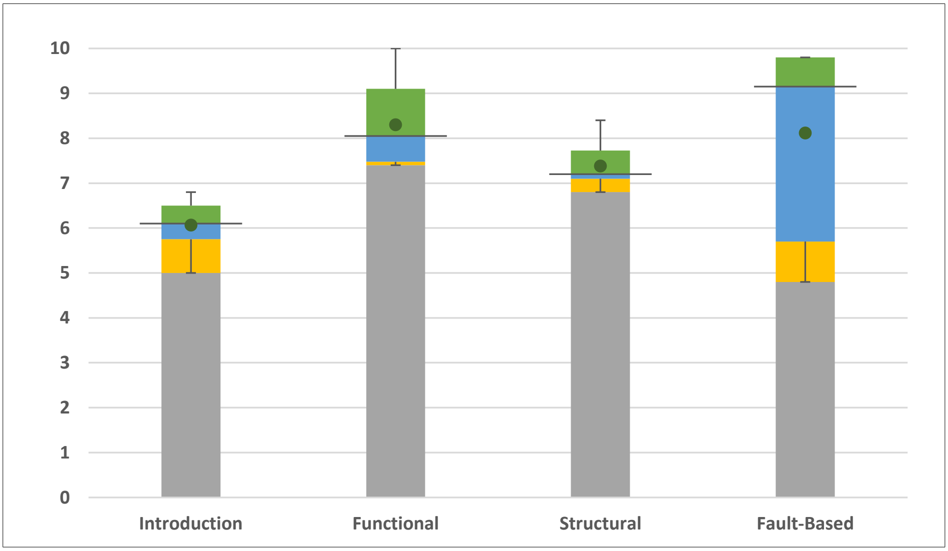

Fig. 5. Boxplot: performance of control group.

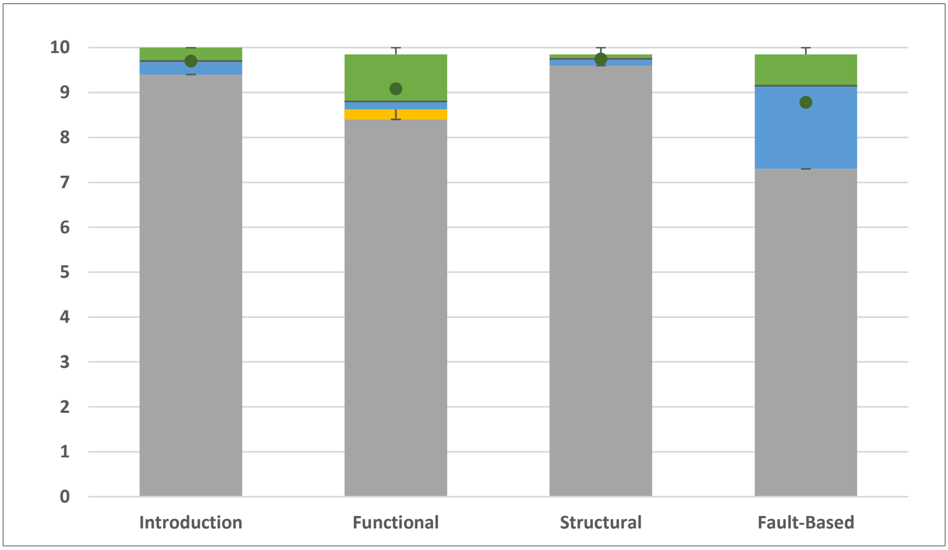

Fig. 6. Boxplot: performance of experimental group

The performance of the experimental group is illustrated in Fig. 6. Note that the data dispersion is smaller than represented in the previous graph. It is possible to observe that participants' performance is more homogeneous in all course topics. As observed in the control group, there are no outliers in the experimental group.

To check the samples normality, Shapiro–Wilks test is also applied. For the control group sample, the *p-value* = 0.822731 was obtained, while for the experimental group sample the *p-value* = 0.136106. In both cases the *p-value* > 0.05 is obtained, so it is possible to consider that the samples are normal.

Assuming the significance value of $\alpha = 0.01$, the values obtained by applying the *t-test* are: *t-value* = 5.01579 and $p = 0.000525$. In this way it is possible to reject the null hypothesis with a two-tailed test at textit p < 0.01.

4.1. *Threats to Validity*

Some threats to validity were identified in this study. We categorized these threats into four categories as proposed by Cook and Campbell (1979). Threats and strategies adopted to mitigate them are discussed below:

- **Conclusion validity**: Conclusions reached through study results were obtained based on statistical test analysis. These tests, if incorrectly conducted, can produce incorrect conclusions. Therefore, the conclusions were based on hypothesis tests and data normalization analysis to mitigate this possible threat.
- **Internal validity**: The conclusions on the causes and their effects can be influenced by other factors that were not considered in the study. Thus, the student's

performance and motivation could be influenced by other factors such as the testing technique covered. We consider the exact topics (four different testing techniques) for both groups to mitigate this threat. In addition, the participants' previous experience with these topics may affect the results. To mitigate the effects of the participants' experience, we use characterization forms to create homogeneous groups with a similar level of knowledge.

- **Construct validity**: Construct validity threats could be added by the selection of the instrument used in the experiment. In this context, we compute the students' performance by the correct answer in multiple-choice tests with the same questions. The motivation was computed through widely used questionnaires from the literature, which were the same for both groups.
- **External validity**: Regarding the generalization of findings obtained, equivalent results may be achieved, provided that it has a selection of the participants done in the academic environment. The experimental planning and execution can be replicated in different software testing topics and teaching-learning processes using SPOCs and distance learning.

5. Conclusions

SPOCs are small and private courses that take place in an online environment. When universities had to be temporarily closed, SPOCs were alternatives so that teaching activities could be carried out. In this study, the main results of the application of an experimental study, aiming to investigate the performance and motivation of students during the application of a SPOC for teaching software testing in times of remote emergency teaching, are presented.

During the study, traditional online teaching methods using a SPOC were applied to two different groups during Verification, Validation, and Software Testing. In summary, the implementation of SPOC went well, although the study was applied in difficult times. The results indicate that the participants in the experimental group had better results, both in performance and motivation. However, a study limitation is how to control the students' learning behavior since, during the application of the experiment, the participants may have suffered from external interference or technical problems, such as difficulty in accessing the Internet.

The application of SPOC with its resources such as forums, video classes, and automated exercises provided distance learning to students less strenuously. These features, along with virtual classroom creation, were crucial to the results obtained in the study. Universities and their professors must be prepared to promote the teaching in emergencies, such as health crises. SPOCs can be an option for face-to-face teaching in emergencies, like that experienced during the COVID-19 pandemic.

Acknowledgments

The authors would like to thank the Brazilian funding agencies: Coordination for the Improvement of Higher Education Personnel – CAPES (Procad) – and National Council for Scientific and Technological Development (CNPq).

References

Alario-Hoyos, C., Estévez-Ayres, I., Delgado Kloos, C., Villena-Román, J. (2017).From MOOCs to SPOCs... and from SPOCs to Flipped Classroom. In: *Data Driven Approaches in Digital Education*, Cham, pp. 347–354. https://doi.org/10.1007/978-3-319-66610-5_25

An, S., Li, W., Hu, J., Ma, L., Xu, J. (2017). Research on the reform of flipped classroom in computer science of university based on SPOC. In: *2017 12th International Conference on Computer Science and Education (ICCSE)*, pp. 621–625. https://doi.org/10.1109/ICCSE.2017.8085567

Basili, V.R., Caldiera, G., Rombach, H.D. (1994). The Goal Question Metric Approach. In: *Encyclopedia of Software Engineering* (Vol. 1). John Wiley & Sons, New Jersey. https://doi.org/10.1002/0471028959.sof142

Chen, C., Zheng, X. (2017). Teaching practise of blended learning on computer general courses for undergraduates. In: *2017 12th International Conference on Computer Science and Education (ICCSE)*, pp. 726–729. https://doi.org/10.1109/ICCSE.2017.8085589

Cheng, M., Zhang, J. (2014). CH-SPOC: A Hybrid Learning Mode and Its Exploration in Zhejiang University. In: *2014 International Conference of Educational Innovation through Technology*, pp. 158–161. https://doi.org/10.1109/EITT.2014.49

Chu, W.C., Hung, S., Chang, C., Ahamed, S.I. (2015). Applying SPOCs for Programming Course to Improve Study Quality. In: *2015 Second International Conference on Trustworthy Systems and Their Applications*, pp. 128–134. https://doi.org/10.1109/TSA.2015.28

Cook, T., Campbell, D. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin, Boston.

Dong, Y., Ang, J., Sun, Z. (2021). Designing Path of SPOC Blended Teaching and Learning Mode in Post-MOOC Era. In: *2021 10th International Conference on Educational and Information Technology (ICEIT)*, pp. 24–28. https://doi.org/10.1109/ICEIT51700.2021.9375582

Filius, R., Kleijn, R., Uijl, S., Prins, F., van Rijen, H., Grobbee, D. (2018). Challenges concerning deep learning in SPOCs. *International Journal of Technology Enhanced Learning*, 10, 111–127. https://doi.org/10.1504/IJTEL.2018.088341

Fox, A. (2013). From MOOCs to SPOCs. *Communications of the ACM*, 56(12), 38–40. https://doi.org/10.1145/2535918

Jong, J.P. (2016). The effect of a blended collaborative learning environment in a small private online course (spoc): a comparison with a lecture course. *Journal of Baltic Science Education*, 15, 194–203. Chap. 194–203. https://doi.org/10.33225/jbse/16.15.194

Kaplan, A.M., Haenlein, M. (2016). Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster. *Business Horizons*, 59(4), 441–450. https://doi.org/10.1016/j.bushor.2016.03.008

Keller, J.M. (2010). The ARCS Model Approach. In: *Motivational Design for Learning and Performance*. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1250-3

Li, X., Gu, C. (2020). Teaching reform of programming basic course based on SPOC blended teaching method. In: *2020 15th International Conference on Computer Science Education (ICCSE)*, pp. 411–415. https://doi.org/10.1109/ICCSE49874.2020.9201802

Liu, A. (2020). Design of Blending Teaching Mode for Software Testing Course. In: *2020 15th International Conference on Computer Science Education (ICCSE)*, pp. 816–821. https://doi.org/10.1109/ICCSE49874.2020.9201740

Lui, R.W.C., Geng, S., Law, K.M.Y. (2017). Project management SPOC with animation. In: *2017 IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, pp. 29–34. https://doi.org/10.1109/TALE.2017.8252299

Martínez–Muñoz, G., Pulido, E. (2015). Using a SPOC to flip the classroom. In: *2015 IEEE Global Engineering Education Conference (EDUCON)*, pp. 431–436. https://doi.org/10.1109/EDUCON.2015.70960070

Mutawa, A.M. (2016). It is time to MOOC and SPOC in the Gulf Region. *Education and Information Technologies*, 1–21. https://doi.org/10.1007/s10639-016-9502-0

Muñoz-Merino, P.J., Rodríguez, E.M., Kloos, C.D., Ruipérez-Valiente, J.A. (2017). Design, Implementation and Evaluation of SPOCs at the Universidad Carlos III de Madrid. *Journal of Universal Computer Science*, 23(2), 167–186. https://doi.org/10.3217/jucs-023-02-0167

Piccioni, M., Estler, C., Meyer, B. (2014). SPOC-supported Introduction to Programming. In: *Proceedings of the 2014 Conference on Innovation and Technology in Computer Science Education*. ITiCSE '14, pp. 3–8. 978-1-4503-2833-3. https://doi.org/10.1145/2591708.2591759

Prates, J.M., Garcia, R., Maldonado, J. (2019). Small Private Online Courses in Computing Learning: evidence, trends and challenges. *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação – SBIE)*, 30(1), 129. https://doi.org/10.5753/cbie.sbie.2019.129

UNESCO (2020). COVID-19 Educational Disruption and Response. Available online (accessed on 30 July 2021): https://en.unesco.org/news/covid-19-educational-disruption-and-response

Wang, M. (2017). The Research and Practice of the Blended Learning Mode in the Construction Courses. In: *Proceedings of the 5th International Conference on Information and Education Technology (ICIET '17)* (Vol. 1). ACM, New York, NY, USA, pp. 58–62. https://doi.org/10.1145/3029387.3029411

Wang, T., Lin, C.-L., Su, Y.-S. (2021). Continuance Intention of University Students and Online Learning during the COVID-19 Pandemic: A Modified Expectation Confirmation Model Perspective. *Sustainability*, 13(8). https://doi.org/10.3390/su13084586

Wohlin, C., Runeson, P., Hst, M., Ohlsson, M.C., Regnell, B., Wessln, A. (2012). *Experimentation in Software Engineering*. Springer Publishing Company, Incorporated, New York, NY.

Zhou, J., Yu, H., Chen, B., Mai, C., Yu, L. (2016). The construction of teaching interaction platform and teaching practice based on SPOC mode. In: *2016 11th International Conference on Computer Science Education (ICCSE)*, pp. 293–298. https://doi.org/10.1109/ICCSE.2016.7581596

Zhou, S., He, Z., Xiong, N., Liu, X. (2019). Research and Application of Mixed Teaching Method of Python Programming Based on SPOC. In: *2019 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pp. 189–193. https://doi.org/10.1109/ICISCAE48440.2019.221615

J.M. Prates is an Adjunct Professor at the State University of Mato Grosso do Sul. He holds degrees in Physics and Computer Science from the São Paulo State University. He received the Ph.D. degree in Computer Science and Computational Mathematics at the Institute of Mathematics and Computer Science (ICMC) University of São Paulo (USP). His research interests are related to the topics of Applied Software Engineering, Computing Applied to Education (Computational Thinking, MOOCs and SPOCs) and Information Visualization.

S.M. Melo received the Ph.D. degree in Computer Science and Computational Mathematics at the Institute of Mathematics and Computer Science (ICMC) University of São Paulo (USP) in 2018. Actually is an Assistant Professor at the Faculty of Exact Sciences and Technology (FACET) Federal University of Grande Dourados (UFGD). Her research interests include computing education, software testing, and experimental software engineering. Contact her at silvanamelo@ufgd.edu.br.

P.H.D. Valle is an Assistant Professor at the Institute of Exact Sciences (ICE) – Federal University of Juiz de Fora (UFJF). He received his BSc (2013) from the Federal University of Goiás (UFG), his Master's of Business Administration Degree in Project Management from the University of São Paulo (ESALQ), and his MSc (2016) and his PhD (2021) in Computer Science from Institute of Mathematics and Computer Science (ICMC) from the University of São Paulo (USP). Currently, his research interests are computer science education, educational games, software architecture, reference architectures, and interoperability.

R.E. Garcia received his Ph.D. degree in Computer Science and Computational Mathematics at the Institute of Mathematics and Computer Science (ICMC) University of São Paulo (USP) in 2006. He is currently Assistant Professor at the São Paulo State University "Júlio de Mesquita Filho", in the Department of Mathematics and Computing of the Faculty of Science and Technology of Presidente Prudente. He is an ad hoc advisor of the São Paulo State Research Support Foundation (FAPESP). He works in the area of computer science, with emphasis on computer methodology and techniques, especially on the following subjects: Software Engineering and Information and Scientific Visualization.

J.C. Maldonado is a Full Professor at the Institute of Mathematics and Computer Science (ICMC) – University of São Paulo (USP), former director of the ICMC, former president of the Brazilian Computer Society – SBC, member of the research network – International Software Engineering Research Network (ISERN, until 2010), and member of the USP Permanent Evaluation Committee. His main topics of interest are software testing, software engineering education, experimental software engineering, critical embedded systems, and teaching environments and methods. He coordinated and participated in dozens of research projects funded by research agencies. He received several awards and honors, including: Scientific Merit Award 2017 from SBC – Brazilian Computer Society, Commendation of Scientific Merit 2008, Jabuti Award 2008 for Exact Sciences, Science and Technology (third place); Highlight Academic 2007 given by ABDI – Brazilian Agency for Industrial Development; and IV National Informatics Award 1991, conferred by SEI, MODDATA S / A, Roberto Marinho Foundation.