Comparison of two online university examination modalities in Didactics of Sciences

Salvador Aljazairi - Universidad de Granada
Inmaculada Salcedo-Bellido - Universidad de Granada
Rocío Barrios-Rodríguez - Universidad de Granada
Pilar Requena - Universidad de Granada

Recepción: 20.04.2022 | Aceptado: 03.05.2022
Correspondencia a través de ORCID: Rocío Barrios Rodríguez 0000-0002-8772-636X

Abstract: Studies about online assessment paying attention to student experiences are scarce. However, the Coronavirus disease 2019 (COVID-19) pandemic has forced universities to adapt very fast to online teaching, which allowed us to analyze the virtual exam modality. We aimed to compare the mark obtained by students in oral and multiple-choice question (MCQ) exams performed online, and to evaluate the students’ satisfaction about them through an anonymous questionnaire. Students enrolled in two subjects of Didactics Sciences area were invited to participate. The exam mark could range from 0 (the worst possible outcome) to 10 (the best one). Results: The participation was high because 90.0% of students took both exams and 87.0% of them filled out the satisfaction questionnaire. Oral exam marks (median=7.0) were significantly higher than the MCQ ones (median=6.3). However, students felt more comfortable and expressed they were more able to show their knowledge with the MCQ than with the oral exam. The main reason why oral exam did not satisfy the students was that “it made them nervous” and “Not doing it” was the most common student answer to improve the oral exam. Conclusion: Although the oral exam benefited students’ mark, it did not satisfy most students.


Comparación de dos modalidades de exámenes virtuales universitarios en Didáctica de Ciencias Experimentales

Resumen: La pandemia por Coronavirus (COVID-19) obligó a las universidades a adaptarse rápidamente a la enseñanza virtual pero los estudios sobre la evaluación no presencial son escasos. Objetivo: comparar la calificación de exámenes orales y de preguntas de opción múltiple (tipo test) virtuales, y evaluar la satisfacción con un cuestionario anónimo, en alumnado del área de Didáctica de las Ciencias Experimentales durante el confinamiento. La calificación de los exámenes oscilaba entre 0 (peor resultado) y 10 (mejor). Resultados: 90% de los alumnos realizaron ambos exámenes y 87% respondieron el cuestionario de satisfacción. La calificación media del examen oral (median=7.0) fue mayor que las del tipo test (median=6.3). Sin embargo, todos los indicadores de satisfacción del examen tipo test obtuvieron mejores puntuaciones que el oral, excepto en la adecuación de la duración del examen. Los estudiantes se sintieron más cómodos y expresaron que eran más capaces de mostrar sus conocimientos con el tipo test. La principal razón por la que los estudiantes no estuvieron satisfechos con el examen oral fue “nerviosismo”. La respuesta más común sobre cómo mejorar el examen oral fue “no hacerlo”. Conclusión: Aunque el examen oral benefició la calificación de los estudiantes, no satisface a la mayoría de los ellos.

Palabra clave: Educación Superior, Evaluación.

Introduction

The pandemic of Coronavirus disease 2019 (COVID-19) and the subsequent city lockdowns and home confinements in many cities around the world in 2020 changed the normal organization at universities. Spain was one of the countries where confinement was more restricted (Hussain, 2020), and students, professors and administrations had to adapt to the new situation very fast (Strielkowski, 2020). As in many countries (Crawford et al., 2020), university face-to-face teaching and examinations were cancelled from March to September 2020 (Gobierno de España, 2020). While teachers...
at University of Granada were required to prioritize, even more than before, the continuous assessment, examinations were yet an important source of evaluation.

Exams should be used as a summative to a variety of assessment tools in online teaching (Meyen, Aust, Bui & Isaacson, 2002; Robles & Braathen, 2002; Rovai, 2000). Alternatives to face-to-face exams using informatics computational technologies can help to assess students confined at home and also, they have benefits for both the students and the teachers (Akimov & Malin, 2020; Rovai, 2000). Comeaux et al. showed some of the benefits of online assessments, among them: i) more efficient management and collection of activities assessed as it is an automatic process in many cases; ii) more opportunities to provide feedback; iii) no restrictions of time and place imposed in face-to-face exams. Moreover, online examination does not affect the performance (score) obtained by students compared to printed exams (Stauffer, Pitlick & Challen, 2020), being an optimal method for evaluation of learning (Boitshwarelo, Reedy & Billany, 2017; Shraim, 2019). However, online examination is not exempt of difficulties: i) it requires infrastructures that some students/faculty members may not have, e.g. computer, broadband internet connection, etc.; ii) universities not used to online teaching may not have the technological capacity to adapt to the required increased use of online platforms, as it happened during the COVID-19 lockdowns(Dill, Fischer, McMurtrie & Supiano, 2020); iii) potential issues of identity security and academic honesty must be taken into account (Akimov & Malin, 2020).

Two of the most commonly used assessment methods are oral and written exams. Rovai et al. (2000) suggested both assessment methods to be taken in online courses. Online written tests are completed by the student in privacy in a given time period, turned in automatically for grading/evaluation (Gharibyan, 2005) and if proctored, they promote identity security and academic honesty (Rovai, 2000). In oral exams an interactive communication exists between the teacher and the student, allowing a deeper evaluation of students’ knowhow(Carnegie, 2015).

Previous studies showed the necessity of more studies about online assessment and examination (Benavides Vázquez & Pedró i García, 2007; Flores Alarcia & Arco Bravo, 2011). Disciplines such as Education might be a useful starting point as these students have provided already endorsement for computerized examinations (Hillier, 2014). Besides, it is necessary to pay careful attention to student experiences to make the learning rich and effective (Sahu, 2020).

For that reason and due to the need of rethinking the evaluation process because online assessment and examination will be more present in the future and will coexist together with face-to-face teaching and examinations; and, in case of future confinements, we carried out this study during the pandemic lockdowns whose main objective was to compare two online exams: oral and multiple-choice questions (MCQ) exams. We analyzed the performance as well as the students’ satisfaction about both modalities.

Methods

This study was carried out in the Faculty of Education, Economy and Technology of Ceuta (University of Granada, Spain) during April 2020.

Sample

Students enrolled in two modules with the same teacher were invited to participate: i) Didactics of Experimental Sciences (3º course of the degree in Primary Education, n=39). In this module students learn the basic principles of Life Sciences as well as how
to design, implement and evaluate practical activities, experiences and teaching resources related to Science and Technology, present in the Primary Education school curriculum. ii) Nutrition and Health Education (1º course of the degree in Early Childhood Education, n=38). In this module, students learn the factors and daily practices related to children’s health, rest, hygiene and activity; as well as nutrition education for children and child development problems related to food.

All students were of legal age and signed an informed consent to participate in the study. They were previously informed about the nature of the study and their volunteer participation, as well as of the exam modalities. Previous to the COVID-19 pandemic, students involved in this study were used to do face-to-face MCQ exams and oral work exhibitions, but they had mostly never performed oral exams. The Vice-Dean for Academic Planning and the Faculty Dean were informed about the study.

**Instruments**

**Modalities of exams**

- **Multiple Choice Questions (MCQ) exam.** Each question had 4-6 answers of which only one correct. Each correct answered added up 1 point, while wrong answers discounted 0.25 points. The score ranged from 0 to 30, but it was eventually transformed into a range 0-10.
- **Oral exam.** Four short random questions were posed out of a pool of 60 questions and the exam lasted 10 minutes. Each correct answered was accounted maximum 2.5 points over 10.

**Student satisfaction questionnaire**

The questionnaire collected the sociodemographic and academic characteristics, and their satisfaction for each type of exam. Sociodemographic data included sex, age, working status and whether they had children. Academic data covered type of evaluation (continuous versus in unique act), average academic grade and previous official examination for that same module. With regards to the exam satisfaction, the following questions were asked for each type of exam: 1) to rate how comfortable they had felt doing the exam (Likert scale, 1-5, being 1 very uncomfortable and 5 very comfortable); 2) if the exam had allowed them to demonstrate their knowledge (yes/no) and why (free text); 3) if they thought that it was a fair system to evaluate their knowledge (yes/no) and why (free text); 4) if the exam duration was adequate (yes/no; if answer was "no", then it should be longer or shorter); 5) how they would improve the exam (free text); and 6) to rate their general satisfaction (Likert scale, 1-10, being 1 very unsatisfied and 10 very satisfied).

**Procedure**

In this study two online modalities of exams were analyzed, which were already part of the assessment planned for that scholar year:

- **Multiple Choice Questions (MCQ) exam.** Five questions were showed in the screen at a time and there was no option to move back to previous questions. There were two types of randomizations: in the questions displayed (30 random questions out of a pool of 120) and in the order of the answers displayed for each question. The exam lasted 30 minutes and was performed through the UGR Moodle-based platform called PRADO-2. It was only available during a specific day and time range.
• **Oral exam.** Each question was posed when the previous question was responded and there was no option to go back to previous questions. The exam took place individually through Google Meet, using a webcam and it was recorded in order to protect the student’s right to review the exam. For the evaluation of each question, a rubric was used that contained: knowledge of the theoretical matter, use of correct scientific vocabulary and scientific expression, didactical applications of knowledge and an appropriate response time.

Both exams covered the theoretical-practical contents of the syllabus. Of note, exams were not the only source of assessment. For instance, didactics resources of teaching were evaluated in an oral essay at the end of the course, but this assessment result was not included in the study as our objective was just to compare the two types of exams. Participants did first the MCQ exam and the same day they did the oral exam. The exams took place in the middle of the semester (not final exams) as a part of the continuous assessment and covered approximately half of the program. For those students that passed the exam the mark accounted for the final score of the module. All the exams and assessments were carried out by the same teacher.

Students filled out an anonymous self-administered online questionnaire after taking both types of exams and before they knew their exam marks.

**Statistical Analysis**

The mean and standard deviation (SD) or median and interquartile range (IQR: 25th and 75th percentiles) of the continuous quantitative variables were calculated. The distribution of absolute and relative frequencies for the qualitative variables was reported. We compared the marks and the student satisfaction between both types of exams. For that, the Wilcoxon signed-rank test for continuous variables, and McNemar test (knowledge and fairness variables) or Fisher exact’ test (duration variable) for qualitative variables were used. Significance was defined at p< 0.05. Analyses and graphs were performed using Stata v.15 (Stata Corp., 2017, College Station, TX, USA) and GraphPad Prism (La Jolla, CA, USA).

**Results**

**Study participants**

In total, 69 (90.0%) students performed both examinations. Of those, 60 (87.0%) students filled out the exam satisfaction questionnaire although some items were not answered by some of them, especially those of free text. The student general characteristics are provided in Table 1. The mean age was 22.7 (SD: 5.1) years and 69.5% were women. Most of them had no children (96.7%) and did not work (85.0%). Most frequently, mean academic grade ranged between 7.01 to 8.00 (out of 10) and a little more than half of the students (62.0%) had never submitted to that module official examination.

**Comparison of exam mark and satisfaction indicators**

Oral exam marks (median=7.0, IQR: 6.0-8.0) were significantly higher than MCQ test marks (median=6.3, IQR: 6.0-7.3, Fig. 1A) (p-value= 0.0318). Students felt more comfortable doing the MCQ than the oral exam (p-value < 0.0001) (Fig. 1B). Overall, student general satisfaction with the MCQ test (median=8, IQR 7.0-9.0) was higher than with the oral exam (median=5, IQR 3.0-7.0) (p-value< 0.0001) (Fig. 1C).
A deeper analysis of the last indicator showed that students of Didactics of Experimental Sciences had greater satisfaction than students of Nutrition and Health Education with both oral and MCQ exams (Fig. 2). However, we found no differences in the oral nor in the MCQ exam satisfaction when we compared students by sex, having or not children, working status, academic grade range, type of evaluation and previous attendance to an official examination (data not shown).

A higher proportion of students (81.0%) declared that the MCQ exam enabled them to show their knowledge compared to the oral one (38.0%, Fig. 3A, p-value < 0.001). Similarly, more students (83.0%) considered the MCQ exam a fair system compared to the oral one (40.0%) to evaluate their knowledge (p-value < 0.001, Fig. 3B). The only indicator that scored worse for the MCQ exam was the adequacy of the duration (59.6% considered adequate the duration for the MCQ exam vs 85.0% for the oral exam, Fig. 3C, p-value= 0.003).
Percentage of students that: a) expressed the exam allowed them to show (YES) or not (NO) their knowledge; b) considered the exam fair to evaluate the knowledge (YES) or not (NO); c) considered the duration of the exam adequate or that it should be shorter/longer. p: p-value, N=60.

Figure 3. Comparison of oral and multiple-choice question (MCQ) exam satisfaction indicators.

Students reasoning about their questionnaire responses

Almost all the students (29/31) that reasoned why the oral exam did not allow them to show their knowledge, blame the nervousness. Similarly, many students considered the oral exam as “not fair” because the nervousness did not allow them to show their real
knowledge. With regards to the students that explained why the oral exam did allow them to show their knowledge (n=18), approximately half of the students (n=10) agreed that the interaction with the teacher promoted a better expression and fluidity of the answers. Those who considered the oral exam fair and gave an explanation why (n=20) provided different reasons that we grouped in three blocks: i) it is as fair as any other evaluation system (40.0%); ii) it allows to show better what you really know as you express yourself (30.0%); iii) it is fair considering the exceptional situation we live (COVID-19 pandemic) and as long as it is combined with other evaluation methods (25.0%).

With regards to the MCQ test and in contrast to the reasons provided for the oral exam, there was not so much consensus in the answers. On the one hand, among those few students that answered why the MCQ test did not allow them to show their knowledge (n=7) or why it was an unfair method (n=5), the most frequent reason provided was that “it was not allowed to move back to previous questions and review the answers”. Other reason alleged was that “the MCQ test did not allow them to show other type of knowledge”. On the other hand, a variety of reasons were provided by those who considered the MCQ test did allow them to demonstrate their knowledge (n=29) and we grouped them in: i) the questions were concise and covered the whole program (31.0%); ii) they have more time to think about the answer to each question (21.0%); iii) the fact of seeing the correct answer written, even if among wrong answers, trigger the memory (17.0%).

Twenty-seven students gave their opinion on how to improve the oral exam. The most common answers were: i) by not doing this type of exam (41.0%) and ii) by practicing this type of examination before (26.0%). Thirty-six students gave their opinion on how to improve the MCQ exam. The most common answers were: i) allowing to move back and forward in the exam to check/change the answers (33.0%) and ii) increasing the time per question (27.0%). Of note, only one student commented that the way to improve the MCQ exam was by not doing it.

Discussion

This study provides relevant data about the performance and satisfaction indicators on two different models of online examinations for an university with classical face-to-face teaching and assessment.

In our study, the performance (measured as mark) was better for the oral than for the MCQ exam, in agreement with previous studies comparing oral vs written exams (Huxham, Campbell & Westwood, 2012; Schickler, Brüstle & Biller, 2015). In fact, 30.0% of students declared that the oral exam allows expressing better what you really know. It has been described that oral exams may give students a better chance to demonstrate their knowledge compared to written tests (Gharibyan, 2005). Moreover, in the classification of competences according to the European educational context, Organic Law 6/2001, of December 21, in universities oral competence appears as an important evaluative element (Roig-Vila et al., 2005). Although MCQ tests are a transparent and economical form of examination that allows covering most of the syllabus contents (Himmelbauer, Koller, Bäwert & Horn, 2019), short open-answer type of questions, as those posed in the oral exam, allow to assess the understanding, interpreting and applying of the existing student knowledge, i.e. competence (Carnegie, 2015; Himmelbauer et al., 2019).

Thus, with previous training for both the teacher (Martín-Cisneros and Aúz-Ramírez, 2007) and the students (Furlan, Alonso-Crespo, Costantini, Díaz-Gutiérrez & Yaryura, 2019), oral exams may be alternatives with proven benefit that should be used more
broadly in courses with continuous assessment (Gharibyan, 2005; Himmelbauer et al., 2019). The disadvantage lies in the amount of work involved in evaluating a large number of candidates by one professor (Himmelbauer et al., 2019).

Despite the finding described above, all the satisfaction indicators in the questionnaires, except exam duration, scored better for the MCQ than for the oral exam: comfort at exam, fairness, capacity to show knowledge and general satisfaction. Not in vain, the most frequent student proposal to improve the oral exam was “not doing it”. Oral examination has been associated with cognitive anxiety (e.g. attention and concentration deficit, negative thoughts, etc.) but not physiological or motor anxiety (Ávila Toscano et al., 2011; Furlan et al., 2019; Iannone and Simpson, 2012). Compared to written tests, students present higher degree of anxiety level with oral exams (Laurin-Barantke, Hoyer, Fehm & Knappe, 2016) and this may add up to the anxiety associated to online examination (Washburn, Herman & Stewart, 2017) and lead to abandoning the course (Furlan et al., 2019). Although we did not measure anxiety in our study, most of the students that declared not feeling comfortable with the oral exam stated feeling nervous and related that nervousness to a poorer performance in the examination. So, even if the performance was actually better, they thought they failed to demonstrate their knowledge with the oral exam, showing a preconceived opinion about this type of examination, probably encouraged by not being used to doing it in regular face-to-face conditions (Akimov and Malin, 2020). Of note, the students did not mention stress or nerves about the pandemic, but just general nervousness about this oral examination. This is in consonance with a previous study showing low levels of anxiety in Higher Education students during the COVID-19 pandemic (Dodd et al., 2021). Although only a minority of students felt comfortable with the oral exam and considered it fair, those who did it, they mentioned that the interaction with the professor allowed a better expression and fluidity of the answers, in agreement with previous publications (Gharibyan, 2005; Iannone and Simpson, 2012; Joughin, 1998).

Written tests are thought to give students privacy and are less intimidating than oral exams (Gharibyan, 2005). According to that, our students did not declare feeling nervous with the MCQ exam and general satisfaction was high. Nevertheless, the two most frequent proposals to improve this examination modality were “allowing to move forward/back in the questions” and “allowing a longer time per question”. However, if implemented, these measures might increase uncertainty about dishonesty. Actually, dealing with the risks of plagiarism and cheating in online exams is more challenging than in the face-to-face ones (Michael and Williams, 2013). This is of special importance due to the lack of experience in online evaluation tools for universities with classical face-to-face teaching and in a period when proctoring tools are being called into question (Flaherty, 2020).

We observed that students of Didactics of Experimental Sciences had a higher satisfaction score than those of Nutrition and Health Education, with both oral and MCQ exams. This might be due to the fact that students of the former were at the 3º course and may have more experience in taking both type of exams. This is in accordance with a previously published study that found that oral exams are more beneficial when students have some experience so they see more clearly the purpose of the proposed activities (Sánchez-Requena, 2018). In fact, the second most frequent proposal to improve the oral exams was “practicing this type of examination before”.

Our study has some limitations. The fact that the participation was volunteer could have led to an overestimation of the positive results as maybe most motivated students enrolled in this study. However, the participation rate in the exam and questionnaire was high, allowing reaching the objectives proposed in this study in a realistic way. The
questionnaire items were not compulsory; therefore, some students did not answer all the questions, especially the open ones. Furthermore, the satisfaction questionnaire used was not previously validated although it allowed to identify possible problems with both type of exams and to reach the objectives proposed in this study.

Our study has also important strengths. We evaluate two types of examinations from two points of view: the outcome and the students’ perception. The answer rate for the satisfaction questionnaire was high (87.0% of students doing the exams). Questionnaires were anonymous to favor sincerity of answers and to minimize the risk of no response. In addition, students filled out the questionnaire before they knew their exam marks, so they were not biased by the result.

Our data showed that student attitudes concerning the examination modalities were not primarily determined by the exam results, while the nervousness related to the oral format due to lack of previous practice played a large role in rejecting this modality. However, the unusual COVID-19 situation has highlighted the necessity of broadening the types of examinations performed at some universities, including oral examination, given its proven benefits and the better marks obtained by the students in our study.

References


Michael, T., & Williams, M. (2013). Student Equity: Discouraging Cheating in Online Courses. Administrative Issues Journal Education Practice and Research, 3(2) 6. https://doi.org/10.5929/2013.3.2.8


