ON THE USE OF MASSIVE EVALUATION TO INCREASE STUDENTS' LEARNING

Desamparados Blazquez, Josep Domenech, Elena de la Poza

Universitat Politècnica de València (SPAIN)
mdeblzso@ade.upv.es, jdomenech@upvnet.upv.es, elpopla@esp.upv.es

Abstract

Assessment is recognized as an essential aspect of student learning, having a crucial effect on it. Indeed, the most important thing lecturers can do for promoting students' learning is to assess their performance. Testing, as a part of the assessment procedure, is more than simply taking a sample of students' learning. It offers the opportunity to provide feedback, which increases motivation and learning if it is prompt and proper. Frequent testing allows students to receive regular feedback.

This article analyzes the effects of assessing students by means of a weekly cumulative testing scheme, which was applied to an intermediate one-semester Microeconomics course. Our experience shows that, if compared with previous years, students are more engaged in the course because they need to attend classes to take the tests. Therefore, a higher ratio of students finished the course. Even more, the experience also shows that massive testing increases the pass rate and overall students' performance.

Keywords: Learning, massive evaluation, feedback.

1 INTRODUCTION

The assessment method used in a course can drive the students' motivation. Motivated students have a good ground for learning, so it is of paramount importance to design assessment methods that properly motivates students to learn. In fact, "what and how students learn depends to a major extent on how they think they will be assessed" [1].

Testing, as a part of the assessment procedure, is more than simply taking a sample of students' learning. It promotes learning, even in the case in which no explicit feedback is provided [2], albeit providing prompt and proper feedback increases motivation and learning to a higher extent [3,4,5]. Frequent testing not only provides students regular feedback, but it also produces valuable information to the assessors when developing the course [6]. However, the literature about the implementation details (such as the exam frequency and content overlap) of a frequent testing scheme is inconclusive [5,7,8]. Hence, more research efforts need to be undertaken.

This paper proposes a new evaluation method in which students' knowledge was weekly put to the test. In addition, the tested contents were cumulative: in each test, students were examined of all explained contents throughout the semester. Testing the same contents repeatedly aims to promote long-lasting learning. This massive evaluation scheme was applied to a Microeconomics module of the first year of a Business degree, which is in line with the new model of education and learning that the European Space for Higher Education (ESHE) requires. This new educational paradigm puts the stress on designing and following continuous evaluation schemes and on fostering a more active role on students. Both objectives are covered by the devised massive evaluation method.

Here, we compare this evaluation scheme with the final examination method, which was applied in previous years. In order to carry out the comparison, we pay attention to some academic performance indicators: the exam and final scores, the withdrawal rate and the pass rate. Results show that the new evaluation method significantly improves the analyzed indicators.

The rest of the paper is organized as follows. Section 2 shows a review of related literature. Section 3 presents the context in which the massive evaluation method was applied. In Section 4 we explain the research method, detailing the specific variables considered in our study as well as the conducted analysis. Section 5 discusses the results of the new evaluation method in contrast to the results of the final examination scheme applied the previous year. Finally, we provide some concluding remarks in Section 6.

2 BACKGROUND

Race [3] and Brown [9] described that the most important deed a teacher can do for students is assessing their performance to help them to learn. Hence, it is of capital importance to devise an effective assessment method which leads students to quality learning and higher performance and engagement on the course.

Traditionally, assessment has been described as “summative” or “formative”. Trotter [10] identifies summative assessment methods as those focusing on the measurement of student achievements up to a given point. In contrast, the main function of formative assessment is providing feedback to students on their performance, as Yorke describes in [6]. However, the distinction between summative and formative assessment is slight, and both perspectives can interact to a greater or lesser extent [11].

Assessment provides information about performance to both students and staff [6, 12] by means of feedback [11]. Feedback plays a major role in education, as it helps students to know their achievement level and staff to reflect on the results of their teaching techniques, allowing both to reshape their learning/teaching methods [6, 13]. Dochy [4] states that assessment can, in fact, be used as a means to consolidate and steer learning.

The essential purpose of feedback is to help students to learn and to improve their performance on a task [11, 14]. Authors like Nicol and Macfarlane-Dick and Crooks [13, 15] show evidence about the effectiveness of feedback in promoting learning. For feedback to be effective, constructive and formative, it must be appropriate, accurate, timely, clear, and focused on student’s personal progress [11,15,16].

Driving a discussion on provided feedback increases its quality [3] and guarantees that students pay attention to that feedback because, as explained by Gibbs and Simpson [11], they participate in a formative dialogue. Consequently, this practice supports their learning process. As feedback yields students to reflect on their progression it must be frequent, being provided when it is still useful for them to identify and correct any deficiencies [12,14,15,16]. Assessing students by continuous assessment is a powerful practice underpinning high-quality learning and teaching; gauging student progress and providing suitable support by means of timely and effective feedback; then, boosting personal reflection and involvement in the learning process [12].

For this reason, applying a frequent testing method allows students to receive regular and prompt feedback. In a context of a frequent assessment methodology, Kling et al. [7] describes that feedback helps students to enhance their learning, leading to boost their academic performance and to increase their engagement. To do so, it is required that feedback is focused in positive comments [13].

3 CONTEXT

The experience is developed at the Universitat Politècnica de València and concerns the assessment methods that were employed during the academic year 2010-11 in Microeconomics II. This course is taught in the second semester of the first year of the Business Administration and Management degree. This program consists of 240 credits, spread over four academic years.

Microeconomics II is the natural continuation to Microeconomics I, a preceding module in which the economic agents –consumers and firms– were introduced as well as their behavior and interactions in the market. Microeconomics II focuses on analyzing the agents’ decision-making process, on reviewing and modeling imperfectly competitive markets and on understanding general equilibrium, market failures and the role of the State.

The course was worth six credits, which corresponds to 60 hours of class and up to 120 hours of autonomous students’ work. It was divided into three parts: lectures, seminars and lab sessions; although this paper only focuses on the lecture part. Lectures had three weekly hours, in which students were divided into two groups, consisting of 52 and 23 students respectively. Besides the lectures, two lab sessions and two seminars were taught, where students were split into smaller groups. Lectures weighted the 85% of the total course score and labs and seminars jointly did the remaining 15%. To pass the course, students were required to get 50% of the score, given they achieve at least 40% of the score of each part and they attend a minimum of 50% of scheduled in-person activities.

Lectures were assessed by weekly hand-written in-class tests. Twelve tests were taken in total, where the average of the best ten tests performance was calculated to obtain the score of this part of the course. These tests had three peculiarities: firstly, they had the duration of 30 minutes, a condition that promoted development of the capability to solve problems quickly and that prevented tiredness; secondly, their accumulative content nature helped students to consolidate the acquired knowledge, as they had to revise the learned material week after week; and finally, two kinds of tests were launched: short quizzes and problem tests, which were taken in weekly alternation. Quizzes had between ten and twelve multiple-choice questions, and problem tests involved one practical problem similar to those explained in class.

Both sections of students took the same test on the same predetermined day. Suggested answers were posted immediately after the tests were taken. Those exam questions that most students did wrong on tests as well as related doubts were tackled in the next day of class. Marks were made available between the same day of the test and two days after the test. The quickly publication of answers and marks sought to provide timely feedback on student's performance, enabling them to identify and amend their weaknesses in time.

4 METHODOLOGY

4.1 Variables

The previously described evaluation scheme was analyzed by comparing massive evaluation academic performance data to previous year data, in which students were assessed using a final examination method. Below, we describe the variables considered in the analysis:

- EXAM_SCORE: It is a continuous variable that represents the fraction of the maximum exam score received by a student in a given exam or test. Hence, it takes values between 0 and 1. As it is defined, students receiving the final examination method have only one exam score, while students receiving the massive evaluation method may have up to 12 scores, depending on their decision.
- FINAL_SCORE: This continuous variable reflects the fraction of the maximum score each student received in the lecture part of the course, applying the corresponding assessment method. That is, this variable takes the same value as EXAM_SCORE for students following the final examination method; but the average of their exam scores (discarding the two worst tests performance) for students in the massive evaluation scheme.
- WITHDRAW: It is a dummy variable that shows whether or not a student finished the course. In the final examination method, we considered that a student did not withdraw if he/she took the final exam. Nevertheless, in the year in which the massive assessment system was applied, a student was considered as withdrawn if he/she did not take 50% of in-class tests.
- PASS: This dummy variable represents whether or not a student passed the course. Following the Spanish scoring system, a student passes the course if he/she obtains at least 50% of the course score.
- YEAR: It is a dummy variable that takes value 0 in the year with the final examination method (2009-2010) and value 1 in the year in which the massive evaluation system was adopted (2010-2011).

4.2 Analysis

To analyze how changing the assessment method affects the academic performance of the students, we firstly perform a descriptive analysis of the variables. Then, to check whether the related differences between both evaluation schemes are statistically significant, a one-way analysis of the variance (ANOVA) is conducted to each considered variable. All data were analyzed using the Statistical Package for the Social Sciences (SPSS) version 16.0 for Windows.

5 RESULTS

Means and standard deviations of the studied variables for each evaluation scheme are reported in Table 1. The descriptive statistics show better results for the massive evaluation method than for the final examination system in all four variables. Below, we analyze in detail these differences.

On examining the differences in the EXAM_SCORE variable, the mean for students following the massive evaluation method was about 13% higher than in the final examination method. Notice that the number of observations is higher in the massive evaluation scheme, due to its massive nature. That is, in the final examination year, each student only has one score; whereas students following the massive evaluation method could have up to 12 marks. The analysis of variance revealed that the differences between the final examination and the massive evaluation methods in the mean EXAM_SCORE are statistically significant. Hence, the presented changes in the evaluation scheme are effective in improving students' exam performance.

The differences in the FINAL_SCORE variable are also similar. Here, the new assessment method provides an increase of about 17 percent points compared to the final examination method. Since the final score in the massive evaluation method is computed as the average of the ten best-scored tests, differences in FINAL_SCORE variable are greater than in the EXAM_SCORE one. As one would expect, the analysis of variance reported significant differences between both treatments.

The means of the WITHDRAW variable shows that the application of the massive evaluation method also contributed to reduce the withdrawal rate from 9% to 5%. However, it is not possible to draw conclusions about the effectiveness of the assessment method in reducing it since the difference between groups is not statistically significant.

Among the studied variables, the pass rate is the one that showed the sharpest improvement. While only 43% of students passed the course using the final examination method, the average of the PASS variable (i.e., the pass rate) reached 93% with the massive evaluation scheme. As expected, the one-way ANOVA revealed that this difference is statistically significant.

6 CONCLUSIONS

This paper sought to evaluate a massive evaluation method, reflecting better results than the final examination method. In the light of the results, we can state that the designed massive evaluation scheme produced several positive effects, which we summarize below.

On the one hand, it promoted a deeper study of the subject as it forced students to revise contents week after week. Also, it increased engagement in the subject, as students need to attend to classes to take the tests.

On the other hand, the withdrawal rate lessened when applying the massive evaluation system, though we cannot conclude if it actually was effective in preventing students from academic withdrawal. Furthermore, this is a difficult point to improve, because withdrawal is highly influenced by students' self-motivation and attitude, and the particular difficulties that they may have for facing problems with the subject (or in general).

In average, the massive evaluation method promoted a better performance with higher scores and a much higher pass rate. This result may be explained because of three reasons: The elimination of the effect of random by doing a higher number of tests per student instead of just the final examination; the increase in students' motivation as they had more opportunities to perform; and, in general, the better knowledge of contents because of the process of continuous revision.

We conclude that applying a massive evaluation scheme is a successful way to steer learning, performance and students' motivation. However, its applicability and effectiveness depends on some other factors, like the difficulty of the subject, its nature (science, social science...) and the number of students in the course. Lecturers should take them into account when designing continuous evaluation methods.

REFERENCES

[1] Biggs, J. and C. Tang (2007). Teaching for quality learning at university (Third Ed). Open University Press.

Table 1 – Descriptive statistics and ANOVA results

	Final examination (YEAR = 0)			Massive evaluation (YEAR = 1)			F value (df = 1)
	N	Mean	S.D.	N	Mean	S.D.	
EXAM_SCORE	133	0.470	0.189	751	0.594	0.243	31.257*
FINAL_SCORE	133	0.470	0.189	71	0.634	0.116	44.203*
WITHDRAW	146	0.090	0.286	75	0.050	0.226	0.885
PASS	133	0.430	0.497	71	0.930	0.258	63.059*

* $P < 0.01$

[2] Dempster, F. N. (1991). Synthesis of research on reviews and tests. Educational Leadership 48, 71-76.

[3] Race, P. (1995). The art of assessing 1. New Academic 4.

[4] Dochy, F. (2008). The Edumetric Quality of New Modes of Assessment: Some Issues and Prospects. Assessment, Learning and Judgement in Higher Education. Dordrecht: Springer Netherlands.

[5] Kuo, T. and A. Simon (2009). How many tests do we really need? College Teaching 57, 156 -160.

[6] Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. Higher Education 45, 477-501.

[7] Kling, N., D. McCorkle, C. Miller, and J. Reardon (2005). The impact of testing frequency on student performance in a marketing course. Journal of Education for Business 81, 67-72.

[8] Casem, M. L. (2006). Active learning is not enough. Journal of College Science Teaching 35.

[9] Brown, S. (2005). Assessment for learning. Learning and Teaching in Higher Education 1, 81-89.

[10] Trotter, E. (2006). Student perceptions of continuous summative assessment. Assessment & Evaluation in Higher Education 31, 505-521.

[11] Brown, G. A., J. Bull, and M. Pendlebury (1997). Assessing Student Learning in Higher Education (1st. Ed.). Routledge.

[12] Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. British Journal of Educational Technology 40, 237-254.

[13] Nicol, D. J. and D. Macfarlane-Dick (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education 31, 199-218.

[14] Chickering, A. W. and Z. F. Gamson (1987). Seven principles for good practice in undergraduate education. American Association for Higher Education Bulletin 39, 3-7.

[15] Crooks, T. J. (1988). The impact of classroom evaluation practices on students. Review of Educational Research 58, 438-481.

[16] Gibbs, G. and C. Simpson (2005). Conditions under which assessment supports students' learning. Learning and Teaching in Higher Education 1 (August 5, 2011), 3-31.