

J.-M. Santos et al.

Influence of continuous assessment / *Influencia de la evaluación continua*

Variation indexes of marks due to continuous assessment. Empirical approach at university / *Índices de variación de la nota debidos a la evaluación continua.*
Contrastación empírica en la enseñanza universitaria

José-Manuel Santos, Esther Ortiz and Salvador Marín

Universidad de Murcia

(Received 12 May 2017; accepted 11 June 2018)

Abstract: The European Higher Education Area (EHEA) has required continuous assessment at university in Spain. The goal of this paper is to analyse how continuous assessment modifies the subject's final mark, to ascertain whether it improves or lowers the final exam score and to determine whether other variables also affect this relationship. To accomplish this, we have created a database with the marks earned by students who took the final exam in the Financial Accounting class at university in the last three academic years. A variation index of the exam score was constructed as a consequence of continuous assessment, and it was analysed both quantitatively and qualitatively. The results obtained from the marks and the index were analysed using contingency tables, ANOVA, regression analyses and a contrast of differences in mean. The main conclusions allow us to state that continuous assessment has an important direct effect on the learning process as well as on the final exam score and the overall class mark, and that other variables also influence them.

Keywords: continuous assessment; summative evaluation; learning process; university; economics

Resumen: La evaluación continua en las notas universitarias es consecuencia de nuestra adaptación al EEES. El objetivo del presente estudio es analizar cómo afecta la evaluación continua a la nota final del alumno, mejorando o empeorando la obtenida en el examen final, y si además otras variables afectan a esta relación. Para ello se ha confeccionado una base de datos con notas de los alumnos presentados al examen final de distintas asignaturas universitarias de Contabilidad Financiera durante los últimos cursos académicos. Se ha construido un índice de variación de la nota del examen como consecuencia de la evaluación continua, y se ha analizado tanto de forma cuantitativa como cualitativa. Las notas y los resultados de este índice se han estudiado con tablas de contingencia, ANOVA, modelos de regresión y un contraste de hipótesis de diferencias de medias. Las principales conclusiones obtenidas permiten probar que efectivamente la evaluación continua tiene un efecto directo sobre el aprendizaje y así sobre la nota obtenida en el examen final y en la asignatura, además que de que influyen otras variables en las notas.

Palabras clave: evaluación continua; evaluación sumativa; proceso de aprendizaje; universidad; ciencias económicas

English version: pp. xx-xx / *Versión en español:* pp. xx-xx

References / *Referencias:* pp. xx-xx

Translated from Spanish / *Traducción del español:* Mary Black

Authors' Address / *Correspondencia con los autores:* José-Manuel Santos, Universidad de Murcia, Economía Financiera y Contabilidad, Murcia, España. E-mail: jmsj1@um.es

The European Higher Education Area (EHEA) has brought about a series of changes in university teaching, and in this study, we shall focus on the effect of continuous assessment on the final exam score, thus changing the overall class mark.

Previous research proves the positive effect of formative assessment on academic performance ([Cano, 2011](#); [Fraile, López, Castejón, & Romero, 2013](#); [Isaksson, 2008](#); [Jareño Cebrián & López García, 2015](#); [Sánchez-Martín, Pascual-Ezama, & Delgado-Jalón, 2017](#)); however, others stress that the cost-benefit ratio of applying formative assessment is not worth it ([Crisp, 2007](#); [Pascual Ezama, Camacho, Ma, Urquía Grande, & Müller, 2011](#); [Poza-Lujan, Calafate, Posadas-Yague, & Cano, 2016](#)), although very few of them use real data with the marks of university students.

In this study, we use real data on university students' marks to analyse the effect of continuous assessment on student learning, which is captured in the score they earn on the final exam, in addition to the potential influence of other variables. The text is organized as follows: first, we highlight the background on the subject; then we describe the sample studied, the methodology used and the hypotheses posited; after that, we describe the results; and then we finish with the conclusions which can be gleaned from this study.

Background

The concept of assessment has evolved in Spain from traditional assessment in the university classroom, the final exam, to the process of continuous assessment in which continuous learning comes to the fore ([Palacios Picos, 2004](#)).

The assessment process is one of the three most important processes that professors carry out, as it is expressed in a mark evaluating the degree to which the learning objectives were met ([Trotter, 2006](#)). It should be set up as a process ([Alcaraz, Fernández, & Sola, 2012](#)) in which the goal is to teach, to carry out a learning process, more than to measure ([Carless, 2007](#)).

Previously, when a single final exam was taken in university studies, the main feature of formative evaluation was not fulfilled: to contribute to students' learning by supplying them with information on how they are progressing (Yorke, 2003). Therefore, in the past, students tended to leave studying for too long (Dowell et al., 2007, p. 4). In addition to limiting the number of tests to just one, it also meant that it was impossible to cater to the vast variety of students to be evaluated (Gibbs & Simpson, 2004; Isaksson, 2008).

This convergence among European universities has led to the implementation of a mixed assessment system (San Martín Gutiérrez, Jiménez Torres, & Jerónimo Sánchez-Beato, 2016, pp. 8–9) in which continuous assessment is applied as part of the final mark, along with a final exam, which weighs more heavily in the final mark of the course. The traditional recourse of assessment via a final exam should only be left for students who do not regularly attend class, or those who have not done the intermediate evaluations called for in continuous assessment (San Martín Gutiérrez et al., 2016, p. 9), because otherwise, the mixed assessment system would be penalizing them.

With the goal of facilitating learning in the course, the number of designs of continuous assessments changes from one academic year to another, and the goal is to design them to cover a wide spectrum of kinds of evaluations, as is desirable for an exhaustive assessment (Crisp, 2007, p. 579). In this sense, Delgado and Oliver (2006) clearly advocate for changing continuous assessment in all teaching periods, for a practical approach to evaluations and for the design and number of evaluations to be a key element in this continuous assessment, which should weigh heavily when calculating the final class mark.

The previous empirical studies have primarily been based on surveys and interviews, as in Hernández (2012), which concludes that the teacher should give feedback on the evaluations they conduct in time and form. Trotter (2006) states that even though continuous assessment is more laborious for the instructor, the cost-benefit ratio is positive in terms of the benefits that students perceive in their learning.

Another line of studies is based on analysing the syllabi of classes, and it concludes that the final exam is still crucial (San Martín Gutiérrez et al., 2016), or that the participative methodology and formative assessment are not used in practice (Castejón & Santos, 2011), which could be a consequence of the fact that adapting educational ideas in another context leads to a failure in their ultimate objectives (Burdett & O'Donnell, 2016).

There are yet other studies which analyse students' academic performance by studying how cooperative learning affects them ([Vallet-Bellmunt, Rivera-Torres, & Vallet-Bellmunt, 2017](#)). [Díaz-Mora, García, and Molina \(2016\)](#) targeted students to ascertain to what extent time use affects their academic performance; they found that academic activities have a positive effect on their learning.

There are studies which underscore the positive effect of formative assessment on academic performance, such as [Cano \(2011\)](#), which found that continuous assessment increases the number of students who pass and improves the marks they earn. [Moro-Egido \(2016\)](#) found that it improves the final mark and also considers written evaluations to be crucial in the final class mark. Similarly, [López-Pastor \(2011, p. 167\)](#) stresses that "these assessment systems seem to be very useful in improving students' academic performance and lowering the failure and drop-out rate in higher education".

[Isaksson \(2008\)](#) found a statistically significant positive relationship between marks and the amount of time students are in a course and therefore are subjected to continuous assessment. [Inda Caro, Álvarez González, and Álvarez Rubio \(2008\)](#) use different types of assessment systems to ascertain which affects students' results the most, and they ultimately found that the more innovative methodologies lead to better results than traditional exams. The study by [\[A\] Fraile et al. \(2003\)](#) also concludes that academic performance improves when the final exam is not used as the sole means of assessment.

In the Financial Accounting course, we can find previous studies such as the one by [Sánchez-Martín et al. \(2017\)](#), which concludes that when students are better informed on the continuous assessment process, their academic results are better. [Pascual Ezama et al. \(2011\)](#) advocate reviewing this assessment methodology, as it does not allow the students with the best scores on the exam to earn a higher mark if they have not earned these marks on the intermediate evaluations, even if there is a positive correlation between the marks on these evaluations and the final exam. [Jareño Cebrián and López García \(2015\)](#) study factors like the results of the courses on Accounting Analysis and Cost Accounting and find that there is a positive relationship between the marks on continuous assessment and the final exam score, although they stress that the final mark in the course is higher with this system than if only the final exam score were borne in mind.

In contrast, [Poza-Lujan et al. \(2016\)](#) quantify the "extra cost" of instruments of using continuous assessment, among other kinds of assessment, and its effect on students' marks; they reach the conclusion that using this kind of assessment has no significant impact on marks. [Crisp \(2007\)](#) cannot conclude that giving students this feedback

improves their marks because there is a series of conditions in university education which prove to be disadvantageous for continuous assessment, such as the number of students in each group and the workload this entails for the instructor ([Yorke, 2003](#)). [Heitink, Van Der Kleij, Veldkamp, and Schildkamp \(2016\)](#) assert that if formative assessment does not have the desired effect on learning it is because it is not implemented correctly, and they thus question whether prerequisites should be required of the instructor, the student, the assessment and the context in order to implement it. They stress the crucial role played by the instructor among the prerequisites analysed through a meta-analysis of previous studies on the subject.

Sample, methodology and hypotheses

We sought to study students' academic results, as captured in their final exam score, and since they may be influenced by the continuous assessment process, as well as whether continuous assessment and marks may be conditioned by other variables. To do so, we developed a database in which we included the marks earned in the last three academic years in the different Financial Accounting classes in the Bachelor's in Business Administration (BBA) and the Bachelor's in Political Science (BPS) at the University of Murcia. This led us to initially have a database of 1,670 marks.

In order to evaluate the students' performance throughout the year, a series of intermediate evaluations were made which have marks and count as a percentage of the final mark in the class. In our case, continuous assessment accounted for **30%** of the final mark. The final exam accounted for the remaining **70%** of the final mark, above the average of **52%** which was found in the study by [San Martín Gutiérrez et al. \(2016\)](#). Likewise, different intermediate evaluations are introduced in the semester to comprise this **30%**.

From the initial database, we eliminated the cases in which the students ultimately did not sit the final exam, as we had to have both marks: continuous assessment and the final exam score, which, in turn, made up the final class mark. Hence, we ultimately had 705 marks, which were broken down into three different expressions:

- The mark on the continuous assessment evaluations;
- The score on the final exam of the course, without taking the continuous assessment percentage into account;

- The final mark in the class, taking the continuous assessment mark into account.

In the initial analysis we calculated a variation rate¹ (which we called *Overall_index*), with the purpose of being able to study in a simple percentage how the score on the exam changed in order to become the final class mark, which is included in expression [1].

$$Overall_index = \left(\frac{\text{Final mark} - \text{Exam score}}{\text{Final mark}} \right) \times 100 \quad [1]$$

The final mark takes into account the percentage from continuous assessment and the exam score, which corresponds solely to the final exam and was traditionally the final mark in the class.

We performed an ANOVA for index values with each of the qualitative variables analysed to check whether there is any pattern and whether the continuous assessment mark adds to the exam score or, on the contrary, subtracts from it depending on the number of partial/intermediate evaluations done, the academic year, the class and the exam period (we included the characteristics of the sample according to these variables in [Table 1](#)). Furthermore, we performed a Tukey post hoc comparison of pairs, considering a p of at least $<.05$ statistically significant.

In the last phase of the empirical analysis, we estimated three different regression models in which the dependent variables were successive: the final exam score (*Exam_mark*) as the learning outcome; the calculated index (*Overall_index*) as the variation that the final exam score undergoes resulting from continuous assessment to get the mark in the class; and finally the continuous assessment mark (*CA_mark*), that is, the mark on the different partial evaluations of learning which are conducted in the class.

The independent variables introduced in each model are the variables among all the possible ones that do not introduce multicollinearity:

- Continuous assessment mark (*CA_mark*). This captures the mark earned on all the partial evaluations which ultimately comprise the continuous assessment mark.
- Number of partial evaluations (*Number_evals*). This captures the number of continuous assessment evaluations in each academic year (see [Table 1](#)).
- Academic year. Based on the academic year when the marks were earned, $n-1$ different dichotomous variables are generated; in this case, since we are examining three academic years, with the goal of avoiding multicollinearity in the model, two variables were introduced: *AY15_16* and *AY16_17*. In both this case and in the other dichotomous variables, the first is taken as the case, such that its effects in the estimation of the model will be captured in the constant.
- Class. In the sample, we included the marks in four different classes ([Table 1](#)), so we got three different dichotomous variables which have the value of 1 when it is the class in question and 0 if not; they are: *Class2351*, *Class2357* and *Class2210*.

- Exam period (*Ex_Per*). There are three exam periods of different exams. Therefore, we introduce two dichotomous variables into the model: *Ex_Per2* and *Ex_Per3*.

Thus, we successively examined the models included in expressions [2], [3] and [4].

$$Exam_mark_i = \beta_0 + \beta_1 CA_mark_i + \beta_2 Number_evals_i + \beta_3 AY15_16_i + \beta_4 AY16_17_i + \beta_5 Class2351_i + \beta_6 Class2357_i + \beta_7 Class2210_i + \beta_8 Ex_Per2_i + \beta_9 Ex_Per3_i + \epsilon_{it} \quad [2]$$

$$Overall_index_i = \beta_0 + \beta_1 Number_evals_i + \beta_2 AY15_16_i + \beta_3 AY16_17_i + \beta_4 Class2351_i + \beta_5 Class2357_i + \beta_6 Class2210_i + \beta_7 Ex_Per2_i + \beta_8 Ex_Per3_i + \epsilon_{it} \quad [3]$$

$$CA_mark_i = \beta_0 + \beta_1 Number_evals_i + \beta_2 AY15_16_i + \beta_3 AY16_17_i + \beta_4 Class2351_i + \beta_5 Class2357_i + \beta_6 Class2210_i + \beta_7 Ex_Per2_i + \beta_8 Ex_Per3_i + \epsilon_{it} \quad [4]$$

We sought to check the hypothesis on whether continuous assessment affects learning and therefore the results obtained on the final exam score (H_1), as well as whether the number of partial evaluations conducted (H_2) influences the continuous assessment and final exam scores, and the academic year (H_3), the class (H_4) and the exam period (H_5). Additionally, we should get coherent results for the index calculated, since the variation in the final exam score as a consequence of continuous assessment will be used as the control results.

In order to conclude the analysis on the influence of continuous assessment on students' learning processes and therefore on the final exam score and the final class mark, we performed a contrast of hypotheses of differences of means in paired samples. Therefore, we applied a *T*-test to three pairs of marks:

Par 1- The continuous assessment mark (*CA_mark*) and the final exam score (*Final_ex_mark*).

Par 2- The continuous assessment mark (*CA_mark*) and the final class mark (*Class_mark*).

Par 3- The final exam score (*Final_ex_mark*) and the final class mark (*Class_mark*).

In any of the three cases, the goal was to check that there are no statistically significant differences in the means of the marks earned and therefore that the same pattern is followed among the different marks. In short, we will have three hypotheses of equal means between each pair of marks.

Results

Analysis of indexes and means

In the first analysis, it seems that having passed the continuous assessment means that the students take the final exam in the first exam period, and the opposite when they have not

passed the successive continuous assessment evaluations (according to the results shown in [Table 2](#)).

Continuous assessment introduces a somewhat noticeable quantitative variation in the final mark compared to the exam score ([Table 3](#)). The final mark may be less than half the exam score, with a variation on the exam score of **-56%**, or the final mark can almost double compared to the exam score if intermediate evaluations are borne in mind (an index of 96.1%). On average, continuous assessment increased the exam score by 10.5%, while it lowered it **-13.2%**.

By isolating the cases in which continuous assessment changes the exam from a passing mark to a failing one, or vice versa, we found 346 cases. Of those with a negative global index — that is, when the exam score is equal to or higher than five but the final mark in the course is lower than five — we found 318 cases that meet this condition. This means that the paradox could arise in which the student passes the exam, but because of the continuous assessment mark they fail the course. The remaining 28 cases have an overall positive index, so continuous assessment is positive for the student, since after failing the final exam they ultimately pass the course.

If we examine to what extent this quantitative mean variation in the exam score affects students by academic year, course, exam period or number of partial evaluations conducted, we find that (included in [Figures 1, 2, 3 and 4](#)):

- When three partial evaluations are conducted, the mean variation in the exam score is 10.2%. When there are two partial evaluations, there is a negative variation, and when four partial evaluations were given, the variation matches the overall mean variation ([Figure 1](#)).

- The greatest variation in the mark when introducing the weight of continuous assessment occurs in the first academic year analysed, as on average it is higher than a **14%** increase ([Figure 2](#)).

- While in courses 2344 and 2357 there is hardly any mean quantitative variation between the exam score and the mark with continuous assessment, this is not true in the other two courses, in which the mean variation reaches almost **11%** in course 2351 and almost the same variation but negative in course 2210 ([Figure 3](#)).

- In the first exam period, the mean variation of the final exam score is **7.42%**, while in the other two it is negative: **-2.3%** in the second exam period and **-4.72%** in the third ([Figure 4](#)). In the first exam period, which is natural for each course, the contribution of

continuous assessment to learning as captured in the final exam is positive, although it is not in the other exam periods.

In all the ANOVA analyses performed, we obtained statistically significant results, as shown in [Tables 4, 5, 6](#) and [7](#). The following are worth highlighting:

- When the number of continuous assessment evaluations increases, it has a positive effect on the final exam score by adding to the exam score (with two evaluations the index is negative, while with three and four it is positive, as shown in the mean in [Table 4](#)). However, in the pair comparisons there is no trend worth noting ([Table 4](#)).

- During two of the three academic years analysed, the continuous assessment mark was positive (in academic years 2014–15 and 2016–17) ([Table 5](#)), while in academic year 2015–16 the learning evaluations throughout the academic year penalized the final mark, and it is also clearly the academic year with the highest number of cases in the sample ([Table 5](#)). These results are also supported when finding statistically significant differences between the first and last academic years studied, and the second (post hoc pair comparisons included in [Table 5](#)).

- In course 2210, the continuous assessment mark subtracts from the student's exam score ([Table 6](#)), even though it is the course with the lowest number of cases in the sample studied. However, in the other courses, these partial evaluations contribute positively to the final mark in the course. The results of the Tukey test show the difference in the contribution of continuous assessment depending on the degree, since statistically significant differences are found between the course in the Bachelor's in Political Science (2210) and the other three courses from the Bachelor's in Business Administration (2344, 2351 and 2357), which also encompass almost all the observations in the sample (data included in [Table 6](#)).

- By studying the exam periods, we found that continuous assessment only contributes positively to the final mark in the first exam period, while in the ensuing periods it becomes a burden, as it subtracts from the exam score and therefore students have to earn a higher score on the exam to pass the course ([Table 7](#) shows a positive mean in the index for the first exam period and negative ones in the second and third periods). This clearly shows that the first exam period is statistically significantly different to the others in terms of the effect of continuous assessment.

Explanatory analysis: the influence of continuous assessment on the mark and on the learning process, as well as the effects of the number of evaluations, the year, the course and the exam period

First, we estimated the regression analyses, as explained in the methodology section, and found statistically significant results in all three cases ([Table 8](#)). In the first two models, the influence of all the independent variables, except in the exam period of the exam, is significant for both the final exam score and the index calculated.

The continuous assessment mark showed a significant, positive relationship with the final exam score, so we can state that our hypothesis 1 (H_1) is accepted, which states that the continuous assessment mark exerts an influence on the final exam score, as shown by [Moro-Egido \(2016\)](#) and as can be clearly deduced from our previous results.

Regarding the second hypothesis, we posited that the number of partial evaluations influenced the other dependent variables, and we can clearly see that in all three models successively there is a statistically significant relationship with the final exam score, the index calculated and the continuous assessment mark, so once again our initial hypothesis (H_2) is accepted. The only caveat in this regard refers to the sign of the relationship found, since while the number of evaluations is positively associated with the final exam score and the continuous assessment mark, it is negative for the coefficient index obtained. This could lead us to argue that while the number of evaluations indeed contributes to learning and to fostering better results on the final exam score and the total mark on continuous assessment, it nonetheless also means making a greater effort on the final exam score in order to improve the course mark, since the partial evaluations may be penalizing it. These results are comparable to those obtained by [Díaz-Mora et al. \(2016\)](#), who conclude that the amount of time students spend on academic activities has positive repercussions on their academic performance, and therefore on their marks. Furthermore, if we further analyse the effect of the number of partial evaluations on the marks, we see that an increase in the number of partial evaluations does indeed imply a higher final mark, which can be explained a priori because part (30%) of the final mark corresponds to the mark earned on continuous assessment, and furthermore an increase in the number of partial evaluations also means an increase in the continuous assessment mark (apart from the results obtained in [Table 4](#), we found significant correlations with $p < .005$ between the number of evaluations and the continuous assessment mark).

The relations obtained with the variables that capture the different academic year are also significant (except for one case), so once again our third hypothesis (H_3) is accepted,

although just as with the number of evaluations, the coefficients are negative for the overall index.

The effect of the courses is statistically significant in all the models, and regardless of the different signs of the coefficients, our fourth initial hypothesis (H₄) can be accepted.

We can clearly not conclude that the exam period has an effect on the relationship between the three dependent variables in the models, which also seems logical for the global index and continuous assessment, since a priori they do not depend on the exam period of the exam for which the student sits. Furthermore, we also found that there is no relationship for the final exam score either (therefore, H₅ cannot be accepted).

To complete the empirical analysis performed, we compared the means of the different marks studied and performed a hypothesis contrast of differences in means in related samples, in this case three pairs of marks, as described in the methodology section. The results of the correlation coefficients ([Table 9](#)) clearly show that the paired variables are all directly related, so when one mark increases so does the other (all the correlations are positive). This result is totally coherent with the design of the marks for pairs 2 and 3, since both the continuous assessment mark and the final exam score are a percentage of the final mark in the course, so any other result would have been impossible. Furthermore, the positive correlation increases as the percentage of the final mark rises (it goes from .551 on continuous assessment to .913 on final exam).

What is truly conclusive is the correlation found for pair 1 of marks, in which there is a direct relationship between the mark earned on continuous assessment and the final exam score, meaning that the basic hypothesis tested in this study is confirmed, namely that continuous assessment truly has a positive effect on student learning and thus on the final exam score, which indirectly affects the final course mark.

In all three cases, there is clearly a statistically significant correlation between the variables, with a level of significance of 95% (the Sig. is 0 in the three pairs of marks < .05) ([Table 9](#)).

We included the results of the hypothesis contrast of different means with paired samples in [Table 10](#). For the three pairs of marks, we have to reject the null hypothesis, since the significance is lower as this is a bilateral contrast in all cases .05 (successively $.001/2 = .0005 < .05$; $0 < .05$; and $.068/2 = .034 < .05$). Therefore, the continuous assessment mark clearly affects the final exam score, as we wanted to test in this study (results obtained for pair 1, since the other two pairs are used as control

variables, because a priori the course note is comprised of the sum of the continuous assessment mark and the final exam score).

Conclusions

After the study performed, we can state that there is a direct influence between continuous assessment and learning and academic performance which is expressed in the final exam score and therefore in the final course mark. These effects are influenced by other variables such as the number of partial evaluations during this process, the academic year and the course. Therefore, engaging in continuous assessment, with different partial evaluations, clearly contributes to greater success on the final exam. And previously, the marks earned on continuous assessment have an effect on the students that sit for the exam in the first exam period, such that if they have passed the intermediate evaluations during the academic year they are considered prepared to sit for the final exam.

Continuous assessment can be used as an indicator that the student has met the learning objectives and is ready for the final evaluation and will therefore pass the class.

Furthermore, even though these intermediate evaluations account for 30% of the final mark, they exert a considerable influence on the score earned on the exam, increasing it an average of 4% and lowering it less than half, so with one exam passed the student can fail the entire course as the result of continuous assessment. However, it is also assumed that greater success on the partial evaluations in continuous assessment will also lead to better results on the final exam score and thus on the total course mark.

We also found that variables like the number of partial evaluations, the academic year and the course may condition the score earned on the final exam and continuous assessment, and the variation in the exam score as a result of continuous assessment to reach the final course mark (the calculated index). In contrast, the exam period in which the student sits for the exam does not seem to have a significant effect on any of these variables, neither the final exam score, nor the mark on continuous assessment, nor the calculated index.

Therefore, continuous assessment does indeed contribute positively to students attaining the learning objectives and passing the final exam and continuous assessment; however, through this same mechanism, students may also earn a lower course grade than they would if only the final exam score were taken into account.

Note

1. This variation rate or index has been used in accounting to attempt to quantify the differences between accounting systems, or to quantitatively measure the harmonization in certain significant periods; conservatism measurement indices have primarily been used, or in a broader sense, indices to measure the degree of comparability, or indices on the concentration of accounting practices (Sánchez et al., 2002).

Índices de variación de la nota debidos a la evaluación continua. Contrastación empírica en la enseñanza universitaria

El Espacio Europeo de Educación Superior (EEES) ha conllevado una serie de cambios en la docencia universitaria, centrándonos en este trabajo en el efecto de la evaluación continua sobre la nota del examen final, modificando así la nota obtenida en total en la asignatura.

Los antecedentes prueban el efecto positivo de la evaluación formativa sobre el rendimiento académico ([Cano, 2011](#); [Fraile, López, Castejón, & Romero, 2013](#); [Isaksson, 2008](#); [Jareño Cebrián & López García, 2015](#); [Sánchez-Martín, Pascual-Ezama, & Delgado-Jalón, 2017](#)); otros en cambio subrayan que no compensa la función coste-beneficio de aplicarla ([Crisp, 2007](#); [Pascual Ezama, Camacho, Ma, Urquía Grande, & Müller, 2011](#); [Poza-Lujan, Calafate, Posadas-Yague, & Cano, 2016](#)), y no muchos de ellos utilizan datos reales con notas de alumnos universitarios.

En este trabajo nos planteamos analizar con datos reales de notas de alumnos universitarios, cuál es el efecto de la evaluación continua en el aprendizaje del alumno, que se plasma en la nota obtenida en el examen final, además de la influencia que puedan tener otras variables. La estructura seguida es la siguiente: en primer lugar, destacamos los antecedentes sobre la materia; seguidamente describimos la muestra estudiada, la metodología y las hipótesis planteadas, a continuación, describimos los resultados obtenidos, para terminar con las conclusiones que nos permite obtener el estudio realizado.

Antecedentes

El concepto de evaluación ha evolucionado en España desde la evaluación tradicional en el aula universitaria: el examen final, hasta el proceso de evaluación continua en el que se prima el aprendizaje continuo ([Palacios Picos, 2004](#)).

El proceso de evaluación es una de las tareas más importantes que realiza el profesor, pues debe plasmar en una calificación el grado de consecución de unos objetivos de aprendizaje ([Trotter, 2006](#)) y se debe configurar como un proceso ([Alcaraz, Fernández,](#)

[& Sola, 2012](#)) en el que se trata de enseñar, llevar a cabo un proceso de aprendizaje, más que medir ([Carless, 2007](#)).

Era anteriormente, al realizar en los estudios universitarios un único examen final, cuando no se cumplía con la característica principal de la evaluación formativa: contribuir al aprendizaje del alumno a través del suministro de información sobre cómo va progresando ([Yorke, 2003](#)) y, por lo tanto, el alumno tendía a dejar el estudio para demasiado tarde ([Dowell et al., 2007](#), p. 4). Además de que el reducir las pruebas a una sola, implica que no se puede dar cabida a la gran diversidad de estudiantes que deben ser evaluados ([Gibbs & Simpson, 2004](#); [Isaksson, 2008](#)).

Esta convergencia entre las universidades europeas ha llevado a la aplicación de un sistema de evaluación mixto ([San Martín Gutiérrez, Jiménez Torres, & Jerónimo Sánchez-Beato, 2016](#), pp. 8-9) en el que se aplica la evaluación continua, que supone un peso de la evaluación final, y una prueba final, con mayor peso en la calificación final de la asignatura. Se debería dejar el recurso de la tradicional evaluación a través del examen final, sólo para aquellos alumnos que no asisten regularmente a clase, o los que no han realizado las pruebas intermedias previstas en la evaluación continua ([San Martín Gutiérrez et al., 2016](#), p. 9), porque si no, el sistema mixto de evaluación les estaría penalizando.

Con el objeto de facilitar el aprendizaje de la asignatura, el número y diseño de las pruebas de evaluación continua se va modificando de un curso académico a otro, y se intentan diseñar cubriendo un amplio espectro de tipos de pruebas, tal y como es deseable para una evaluación exhaustiva ([Crisp, 2007](#), p. 579). En este sentido Delgado y Oliver ([2006](#)) abogan claramente por la conveniencia de cambiar las pruebas de evaluación continua todos los períodos lectivos, por el enfoque práctico de las pruebas realizadas, y por el hecho de que el diseño y número de pruebas es un elemento clave de esta evaluación continua, que deben tener un **valor** importante para calcular la nota final de la asignatura.

Los antecedentes mediante estudios empíricos se han basado principalmente en la realización de encuestas y entrevistas, como [Hernández \(2012\)](#) que concluye que el docente debe dar el retorno, sobre las pruebas que se vayan realizando, en tiempo y forma. [Trotter \(2006\)](#) afirma que, aunque la evaluación continua es más trabajosa para el docente, la función coste-beneficio resulta positiva, según los beneficios en cuanto a aprendizaje que perciben los alumnos.

Otra línea de estudios, se basa en el análisis de las guías docentes de las asignaturas, concluyendo que el examen final sigue siendo clave ([San Martín Gutiérrez et al., 2016](#)), o que en la práctica no se utiliza la metodología participativa ni la evaluación formativa ([Castejón & Santos, 2011](#)), lo que puede ser consecuencia de que adaptar ideas educativas en otro contexto lleva a fallar en sus objetivos finales ([Burdett & O'Donnell, 2016](#)).

Y también existen otros trabajos que sí que analizan rendimientos académicos de los alumnos, pero estudiando cómo influye en ellos el aprendizaje cooperativo ([Vallet-Bellmunt, Rivera-Torres, & Vallet-Bellmunt, 2017](#)), o Díaz-Mora, García, y Molina (2016) que se dirigen a los alumnos para comprobar en qué medida el uso de su tiempo incide en su rendimiento académico, comprobando que las actividades académicas inciden positivamente en su aprendizaje.

Existen estudios que subrayan el efecto positivo de la evaluación formativa sobre el rendimiento académico, como [Cano \(2011\)](#), que determina que la evaluación continua aumenta el número de aprobados y mejora las notas obtenidas. [Moro-Egido \(2016\)](#) prueba que mejora la calificación final y además plantea que la tipología de pruebas escritas es determinante para la calificación final de la asignatura. En este mismo sentido, [López-Pastor \(2011, p. 167\)](#), resaltan que «estos sistemas de evaluación parecen ser muy útiles en la mejora del rendimiento académico del alumnado y en la disminución de las tasas de fracaso y abandono en los estudios superiores».

[Isaksson \(2008\)](#) sí que obtiene una relación positiva estadísticamente significativa, entre las notas y el tiempo que los estudiantes están en el curso, y por lo tanto se someten a evaluación continua. Inda, Álvarez, y Álvarez (2008) utilizan distintos tipos de sistemas de evaluación para ver cuál es el que más incide en los resultados de los alumnos, destacando finalmente que las metodologías más innovadoras, frente al examen tradicional, llevan a mejores resultados de los alumnos. El trabajo de [Fraile et al. \(2003\)](#) también concluye que el rendimiento académico mejora cuando no se usa como única vía el examen final.

Para el caso de la asignatura de Contabilidad Financiera, encontramos antecedentes como el de [Sánchez-Martín et al. \(2017\)](#), en el que concluyen que cuando los estudiantes están mejor informados sobre este proceso de evaluación continua obtienen mejores resultados académicos. [Pascual Ezama et al. \(2011\)](#) abogan por revisar este método de evaluación, pues no permite a los estudiantes con mejores notas en el examen final obtener una mejor calificación si no lo han demostrado en las notas de las pruebas intermedias, aunque existe correlación positiva entre las notas obtenidas en estas pruebas y la nota del

examen final. Jareño y López (2015) estudian, entre otros, los resultados de las asignaturas de Análisis Contable y Contabilidad de Costes, y obtienen que existe correlación positiva entre las calificaciones de la evaluación continua y la nota obtenida en el examen final, pero destacan que la nota final en la asignatura es mayor con este sistema, que si sólo se hubiera tenido en cuenta la nota del examen final.

En cambio, Poza-Lujan et al. (2016) cuantifican el **sobrecoste** que supone para el docente la evaluación continua, entre otros tipos de evaluación, y su efecto en la nota de los estudiantes, llegando a la conclusión de que utilizar este tipo de evaluación no tiene un impacto significativo en las notas. Crisp (2007) no puede concluir que el hecho de dar ese **feedback** a los alumnos mejore las calificaciones, porque existen una serie de condiciones de la educación universitaria que suponen inconvenientes para la evaluación continua, como el número de alumnos en cada grupo, o la carga de trabajo que supone para el docente (Yorke, 2003). Heitink et al. (2016) sostienen que, si la evaluación formativa no tiene el efecto deseable sobre el aprendizaje, es porque no está correctamente implementada, y para ello se plantean si tienen que existir unos requisitos previos con respecto al profesor, al estudiante, la evaluación y el contexto, de cara a esta implantación. Resaltan el papel crucial que juega el profesor entre los requisitos previos analizados a través de meta análisis de estudios previos sobre la materia.

Muestra, metodología e hipótesis

Pretendemos estudiar los resultados académicos de los alumnos, que se concretan en la nota del examen final, y cómo pueden estar influenciados por el proceso de evaluación continua, así como a su vez si la evaluación continua y las notas pueden verse condicionadas por otras variables. Para ello hemos elaborado una base de datos en la que incluimos las notas obtenidas en los tres últimos cursos académicos, en las distintas asignaturas de Contabilidad Financiera en el Grado en Administración y Dirección de Empresas (ADE) y en el de Ciencias Políticas (CC Políticas) de la Universidad de Murcia. Lo que nos lleva a contar inicialmente con una base de datos de 1,670 notas.

Para evaluar el rendimiento del alumno a lo largo del curso se realizan una serie de pruebas de evaluación intermedias, que cuentan con una calificación y suponen un porcentaje de la nota final. En nuestro caso la evaluación continua supone el **30%** de la calificación final. El examen final supone el restante **70%**, por encima de la media de un

52%, que se obtiene en el trabajo de [San Martín Gutiérrez et al. \(2016\)](#). A su vez para calificar este 30% se introducen distintas pruebas intermedias en el cuatrimestre.

De la base de datos inicial eliminamos aquellos casos en los que el alumno no se presenta finalmente al examen final, pues pretendemos contar con ambas notas: la nota de la evaluación continua, y la nota del examen final, que a su vez luego se concretarán en la nota final de la asignatura. De ahí que finalmente contamos con 705 notas, que su vez se desglosan en tres expresiones distintas:

- La nota de las pruebas de evaluación continua;
- La nota del examen de la asignatura, sin tener en cuenta el porcentaje de evaluación continua;
- La nota final de la asignatura teniendo en cuenta la evaluación continua.

En el análisis inicial hemos calculado una tasa de variación¹ [link nota 1](#) (que denominaremos *Indice_global*), con el objeto de poder estudiar de una forma sencilla y en porcentaje, cómo ha cambiado la nota del examen, para llegar a ser la nota final de la asignatura, lo que se incluye en la expresión [1].

$$Indice_global = \frac{(Nota\ final - Nota\ examen)}{Nota\ final} \times 100 \quad [1]$$

Siendo la nota final aquella en la que se tiene en cuenta el porcentaje correspondiente a evaluación continua, y la nota del examen la que corresponde únicamente al examen final y que tradicionalmente sería la nota final de la asignatura.

Hemos realizado un ANOVA para los valores del índice con cada una de las variables cualitativas analizadas, para comprobar si existe algún patrón en el mismo, y si la nota de la evaluación continua suma a la nota del examen, o por el contrario le resta, dependiendo del número de pruebas parciales/intermedias realizadas, el curso académico, la asignatura, o la convocatoria (las características de la muestra en función de estas variables las incluimos en la [Tabla 1](#)). Además, realizamos una comparación por pares post hoc a través de Tukey considerando significatividad estadística con una p al menos $<.05$.

En la última fase del análisis empírico hemos estimado tres diferentes modelos de regresión, en los que las variables dependientes serán sucesivamente: la nota del examen final (*nota_examen*), como resultado del aprendizaje; el índice calculado (*Indice_global*), como variación que sufre la nota del examen final para obtener la nota de la asignatura consecuencia de la evaluación continua; y finalmente la nota de la evaluación continua (*Nota_eva_cont*), es decir, la nota de las distintas pruebas parciales que se van desarrollando para el aprendizaje durante el curso.

Las variables independientes introducidas en cada modelo serán las que no introducen multicolinealidad de entre todas las posibles:

- Nota de la evaluación continua (*Nota_eva_cont*). Recoge la nota obtenida en todas las pruebas parciales que al final dan lugar a la nota de la evaluación continua.
- Número de pruebas parciales (*Número_pruebas*). Recoge el número de pruebas de evaluación continua realizadas en cada curso académico (ver [Tabla 1](#)).
- Curso académico; A partir del curso académico del que se hayan obtenido las notas se generan n-1 variables dicotómicas distintas, en este caso al tratarse de tres cursos académicos, con el objeto de evitar multicolinealidad en el modelo se introducirán dos variables: *Curso15_16* y *Curso16_17*. Tanto en este caso como en el resto de las variables dicotómicas tomaremos como base la primera, de manera que sus efectos en la estimación del modelo quedarán recogidos en la constante.
- Asignatura; En la muestra incluimos las notas de cuatro asignaturas distintas ([Tabla 1](#)), por lo tanto, obtendremos tres variables dicotómicas distintas, que tomarán valor 1 cuando se trate de la asignatura en cuestión y 0 en el caso contrario, siendo las siguientes: *Asign2351*, *Asign2357*, y *Asign2210*.
- Convocatoria (*Conv*); Existen tres convocatorias de exámenes distintas. Por ello introducimos en el modelo dos variables dicotómicas: *Conv2* y *Conv3*.

Por lo tanto planteamos sucesivamente los modelos incluidos en las expresiones [2], [3] y [4].

$$\begin{aligned} \text{Nota_examen}_i = & \beta_0 + \beta_1 \text{Nota_eva_cont}_i + \beta_2 \text{Número_pruebas}_i + \beta_3 \text{Curso15_16}_i \\ & + \beta_4 \text{Curso16_17}_i + \beta_5 \text{Asign2351}_i + \beta_6 \text{Asign2357}_i + \beta_7 \text{Asign2210}_i + \beta_8 \text{Conv2}_i + \beta_9 \text{Conv3}_i \\ & + \varepsilon_{it} \quad [2] \end{aligned}$$

$$\begin{aligned} \text{Índice_global}_i = & \beta_0 + \beta_1 \text{Número_pruebas}_i + \beta_2 \text{Curso15_16}_i + \beta_3 \text{Curso16_17}_i + \beta_4 \\ & \text{Asign2351}_i + \beta_5 \text{Asign2357}_i + \beta_6 \text{Asign2210}_i + \beta_7 \text{Conv2}_i + \beta_8 \text{Conv3}_i + \varepsilon_{it} \quad [3] \end{aligned}$$

$$\begin{aligned} \text{Nota_eva_cont}_i = & \beta_0 + \beta_1 \text{Número_pruebas}_i + \beta_2 \text{Curso15_16}_i + \beta_3 \text{Curso16_17}_i + \beta_4 \\ & \text{Asign2351}_i + \beta_5 \text{Asign2357}_i + \beta_6 \text{Asign2210}_i + \beta_7 \text{Conv2}_i + \beta_8 \text{Conv3}_i + \varepsilon_{it} \quad [4] \end{aligned}$$

Pretendemos comprobar la hipótesis relativa a sí la evaluación continua afecta al aprendizaje y por lo tanto a los resultados obtenidos en la nota del examen final (H_1), así como si el número de pruebas parciales realizadas (H_2) influye sobre la nota de la evaluación continua y la del examen final, y también el curso académico (H_3), la asignatura de la que se trate (H_4) o la convocatoria del examen (H_5). Por añadidura nos deben salir resultados coherentes para el índice calculado, ya que es la variación de la nota del examen consecuencia de la evaluación continua, que nos servirán como resultados de control.

Con el objeto de concluir el análisis sobre la influencia de la evaluación continua en el proceso de aprendizaje del alumno, y por lo tanto en la nota del examen final y en la

nota final de la asignatura, realizamos un contraste de hipótesis de diferencias de medias en muestras pareadas, o emparejadas. De manera que aplicamos una prueba T a tres pares de notas:

Par 1- La nota de la evaluación continua ($Nota_eva_cont$) y la nota del examen final ($Nota_exa_fin$).

Par 2- La nota de la evaluación continua ($Nota_eva_cont$) y la nota final de asignatura ($Nota_fin_asig$).

Par 3- La nota del examen final ($Nota_exa_fin$) y la nota final de la asignatura ($Nota_fin_asig$).

En cualquiera de los tres casos se trata de comprobar que no existen diferencias estadísticamente significativas en las medias de las notas obtenidas, y por lo tanto, que se sigue el mismo patrón entre las distintas notas. En definitiva, tendremos tres hipótesis de medias iguales entre cada par de notas.

Resultados

Análisis de los índices y medias

En un primer análisis parece que el haber superado la evaluación continua hace que los alumnos se presenten en primera convocatoria al examen final, y al contrario cuando no se han superado las sucesivas pruebas de evaluación continua (según los resultados obtenidos en la [Tabla 2](#)).

La evaluación continua introduce una variación cuantitativa no despreciable en la nota final con respecto a la del examen ([Tabla 3](#)). La nota final puede quedarse en algo menos de la mitad de la del examen, con una variación de la nota del examen del **-56%**, o la nota final casi llegar a doblarse con respecto a la del examen al tener en cuenta las pruebas intermedias (un índice del **96.1%**). Por término medio la evaluación continua incrementa la nota del examen en un **10.5%**, mientras que la disminuye en un **-13.2%**.

Aislando los casos en los que la evaluación continua cambia de un aprobado en el examen a un suspenso o, al contrario, nos encontramos con 346 casos. De ellos los que tienen un índice global negativo, es decir, la nota del examen es igual o superior a cinco, pero en cambio la nota final de la asignatura es inferior a cinco, nos encontramos con 318 casos que cumplen estas condiciones. Lo que supone que se puede dar la paradoja de que el alumno supera el examen, pero en cambio debido a la nota de la evaluación continua, suspende la asignatura. Los restantes 28 casos tienen un índice global positivo, con lo que

al alumno le está resultando positiva la evaluación continua, pues habiendo suspendido el examen final finalmente logra aprobar la asignatura.

Si observamos en qué medida la media de esta variación cuantitativa en la nota del examen afecta por curso académico, asignatura, convocatoria o según el número de pruebas parciales realizadas, comprobamos que (incluido en las [Figuras 1, 2, 3 y 4](#)):

- Cuando se realizan tres pruebas parciales la variación media en la nota del examen es de un 10.20%. En el caso de dos pruebas se produce una variación negativa, y cuando se han puesto cuatro pruebas la variación coincide con la variación media global ([Figura 1](#)).

- La mayor variación en la nota al introducir la ponderación de la evaluación continua se produce en el primer curso académico analizado, superando en término medio un 14% de incremento ([Figura 2](#)).

- Mientras que en las asignaturas 2344 y 2357 no hay apenas variación media cuantitativa entre la nota del examen y la nota con evaluación continua, no ocurre así en las otras dos asignaturas, en las que la variación media casi alcanza el 11% en la asignatura 2351 y casi la misma variación, pero con signo negativo en la 2210 ([Figura 3](#)).

- En la primera convocatoria la variación media de la nota del examen final es de un 7.42%, mientras en las otras dos tiene signo negativo, siendo de -2.3% en la segunda convocatoria y de -4.72% en la tercera ([Figura 4](#)). En la primera convocatoria, que resulta natural para cada asignatura, la contribución de la evaluación continua al aprendizaje que se plasma en el examen final es positiva, y no así en el resto de convocatorias.

En todos los análisis ANOVA realizados hemos obtenido resultados estadísticamente significativos, tal y como se recoge en las [Tablas 4, 5, 6 y 7](#), entre los que podemos resaltar los siguientes:

- Cuando el número de pruebas de evaluación continua aumenta tienen un efecto positivo sobre la nota del examen final, sumando a la nota del examen (con dos pruebas el índice es negativo, mientras que con tres y cuatro es positivo, como se aprecia en la media de la [Tabla 4](#)). Aunque en la comparación por parejas no existe una tendencia a destacar ([Tabla 4](#)).

- Durante dos de los tres cursos académicos analizados la nota de la evaluación continua ha sido positiva (en los cursos 2014-15 y en el 2016-17) ([Tabla 5](#)), mientras que en el curso 2015-16 las pruebas de aprendizaje a lo largo del curso han penalizado a la nota final, siendo además claramente el curso con mayor número de casos en nuestra muestra ([Tabla 5](#)). Resultados que además se refrendan al obtener diferencias

estadísticamente significativas entre el primer y último curso estudiado, y el segundo (comparaciones por pares post hoc incluidos en la [Tabla 5](#)).

- En la asignatura 2210 la nota de evaluación continua resta a la nota del examen del alumno ([Tabla 6](#)), aunque es la asignatura con un menor número de casos en la muestra estudiada. Mientras que en el resto de asignaturas estas pruebas parciales contribuyen de forma positiva a la nota final de la asignatura. Los resultados de Tukey muestran la diferencia en la contribución de la evaluación continua dependiendo del Grado, pues se obtienen diferencias estadísticamente significativas entre la asignatura del Grado en Ciencias Políticas (2210) y las otras tres asignaturas del Grado en Administración y Dirección de Empresas (2344, 2351 y 2357), que además aglutinan la casi totalidad de las observaciones de la muestra (datos incluidos en la [Tabla 6](#)).

- A partir del estudio de las convocatorias se obtiene que sólo contribuye de forma positiva a la nota final la evaluación continua en la primera convocatoria, mientras que en las siguientes se convierte en un lastre, pues resta a la nota del examen y por lo tanto para aprobar la asignatura el alumno debe obtener mejor calificación en el examen (en la [Tabla 7](#) se obtiene media positiva del índice en la primera convocatoria y negativas en la segunda y tercera). Obteniéndose claramente que la primera convocatoria es diferente de forma estadísticamente significativa con respecto a las siguientes, en cuanto al efecto de la evaluación continua.

Análisis explicativo: la influencia de la evaluación continua en la nota y en el proceso de aprendizaje, así como los efectos del número de pruebas, el curso, la asignatura y la convocatoria

En primer lugar estimamos los modelos de regresión, tal y como se ha explicado en el apartado de metodología, obteniendo unos resultados estadísticamente significativos en los tres casos ([Tabla 8](#)). En los dos primeros modelos planteados, tanto para la nota del examen final, como para el índice calculado, la influencia de todas las variables independientes es significativa, excepto la convocatoria del examen.

La nota de la evaluación continua obtenemos que tiene una relación significativa y positiva con la nota del examen final, con lo que podemos afirmar que aceptamos nuestra hipótesis 1 (H_1), según la cual existe una influencia de la nota de la evaluación continua en la del examen final, tal y como también se demuestra por [Moro-Egido \(2016\)](#) y se podía deducir claramente de nuestros resultados anteriores.

Sobre la segunda hipótesis que planteamos, acerca de la existencia de influencia del número de pruebas parciales realizadas en las otras variables dependientes, claramente vemos que en los tres modelos planteados se obtiene una relación estadísticamente significativa, sucesivamente, con la nota del examen final, el índice calculado, y la nota de la evaluación continua, con lo que de nuevo se acepta nuestra hipótesis de partida (H_2). La única puntualización al respecto se refiere al signo de la relación obtenida, ya que mientras que el número de pruebas se vincula de forma positiva con la nota del examen final y la de la evaluación continua, para el índice el coeficiente obtenido es negativo. Esto nos podría llevar a argumentar que mientras que efectivamente un mayor número de pruebas contribuye al aprendizaje y a fomentar unos mejores resultados en la nota del examen final y en la nota total de la evaluación continua, en cambio conlleva tener que hacer un mayor esfuerzo en la nota del examen final para poder mejorar la nota de la asignatura, pues más pruebas parciales la están penalizando. Podemos asemejar estos resultados a los obtenidos por Díaz-Mora, García, y Molina (2016), cuando concluyen que el tiempo que los estudiantes dedican a actividades académicas repercute de forma positiva en su rendimiento académico, y por lo tanto en su nota. Además, ahondando más en el análisis del efecto del número de pruebas parciales en las notas, obtenemos que efectivamente un aumento de las mismas implica una nota final más alta, lo que viene explicado a priori porque una parte (el 30%) de la nota final corresponde a la nota obtenida de la evaluación continua, pero es que además un aumento en el número de pruebas parciales también implica un aumento en la nota de la evaluación continua (aparte de los resultados obtenidos en la [Tabla 4](#) obtenemos correlaciones significativas con $p < .005$ entre el número de pruebas y la nota de la evaluación continua).

Las relaciones obtenidas con las variables que recogen el distinto curso académico también resultan significativas (excepto un caso puntual), con lo que de nuevo podemos aceptar nuestra tercera hipótesis (H_3), aunque como sucede con el número de pruebas, para el índice global los coeficientes obtenidos son negativos.

El efecto de las distintas asignaturas es estadísticamente significativo en todos los modelos, y con independencia de los distintos signos en los coeficientes, sí que podemos aceptar nuestra cuarta hipótesis de partida (H_4).

Claramente no podemos concluir que exista un efecto de la convocatoria en la relación existente entre las tres variables dependientes de los modelos planteados, lo cual además parece lógico para el caso del índice global y la evaluación continua, pues a priori no dependen de la convocatoria del examen a la que se presente el alumno, pero es que

además también obtenemos que no existe relación tampoco para la nota del examen final (no se puede aceptar H_5).

Para terminar el análisis empírico realizado comparamos las medias de las distintas notas estudiadas, realizando un contraste de hipótesis de diferencias de medias en muestras relacionadas, en este caso, como ya hemos descrito en el apartado de metodología, se trataría de tres pares de notas. Los resultados obtenidos de los coeficientes de correlación ([Tabla 9](#)) claramente nos dicen que las variables emparejadas están todas relacionadas en sentido directo, cuando una nota crece la otra también (todas las correlaciones son positivas). Este resultado es totalmente coherente con el diseño de las notas para las parejas 2 y 3, pues tanto la nota de la evaluación continua, como la nota del examen final, constituyen un porcentaje de la nota final de la asignatura, con lo que la obtención de otro resultado no hubiera sido posible. Además, la correlación positiva aumenta conforme aumenta el porcentaje de contribución a la nota final (pasa de un .551 de la evaluación continua a un .913 del examen final).

Lo verdaderamente concluyente es la correlación obtenida para la pareja 1 de notas, según la cual existe una relación directa entre la nota obtenida en la evaluación continua y la nota del examen final, o lo que es lo mismo, se vuelve a refrendar la hipótesis básica a contrastar en este trabajo, que verdaderamente la evaluación continua incide de forma positiva en el aprendizaje del alumno y así en la nota del examen final, lo que de forma indirecta repercute en la nota final de la asignatura.

Claramente en los tres casos se obtiene una correlación significativamente estadística entre las variables con un nivel de significatividad del 95% (la Sig. es de 0 en las tres parejas de notas $<.05$) ([Tabla 9](#)).

En la [Tabla 10](#) incluimos los resultados del contraste de hipótesis de diferencias de medias con muestras emparejadas. Para los tres pares de notas se debe rechazar la hipótesis nula, ya que al ser un contraste bilateral en todos los casos la significatividad es menor a .05 (sucesivamente $.001/2 = .0005 <.05$; $0 <.05$; y $.068/2 = .034 <.05$). Por lo tanto, claramente la nota de la evaluación continua afecta a la nota del examen final, tal y como queríamos contrastar en este trabajo (resultados obtenidos para el par 1, ya que los otros dos pares sirven de variables de control, pues a priori la nota de la asignatura está compuesta por la suma de la nota de la evaluación continua y del examen final).

Conclusiones

Tras el estudio realizado podemos afirmar que existe una influencia directa entre la

evaluación continua y el aprendizaje y rendimiento académico, que se concreta en la nota del examen final, y por lo tanto en la nota final de la asignatura. Estos efectos vienen influidos por otras variables como el número de pruebas parciales que se realicen en este proceso, el curso académico, o la asignatura de la que se trate. Por lo tanto, claramente, el realizar un proceso de evaluación continua, con diferentes pruebas parciales, contribuye a que haya mayor éxito en el examen final. Y previamente las notas obtenidas en la evaluación continua tienen un efecto sobre los alumnos que se presentan o no a la primera convocatoria del examen, de manera que si han superado las pruebas intermedias durante el curso se consideran preparados para concurrir al examen final.

La evaluación continua se puede configurar como un indicador de que se han superado los objetivos del aprendizaje, y el alumno está preparado para la prueba final, y por lo tanto superará la asignatura.

Además, aunque estas pruebas intermedias supongan el 30% de la nota final, sí que afectan cuantitativamente de forma no despreciable a la nota obtenida en el examen, incrementándola en términos medios más de un 4%, y llegando a reducirla por debajo de la mitad, con lo que, con un examen aprobado, el alumno puede suspender la asignatura como consecuencia de la evaluación continua. Aunque también se supone que un mayor éxito en las pruebas parciales de evaluación continua conllevaría unos mejores resultados en la nota del examen final, y así, en la calificación total de la asignatura.

También obtenemos que variables como el número de pruebas parciales, el curso académico, y la asignatura, pueden condicionar la nota obtenida en el examen final, en la evaluación continua, y por añadidura la variación que sufre la nota del examen como consecuencia de la evaluación continua para llegar a la calificación final de la asignatura (el índice calculado). En cambio, la convocatoria de examen no obtenemos que afecte de forma significativa a ninguna de estas variables: ni a la nota del examen final, ni a la nota de la evaluación continua, ni al índice calculado.

De manera que sí efectivamente con la evaluación continua se puede contribuir de forma positiva a que los alumnos consigan los objetivos del aprendizaje, superando el examen final y la evaluación continua, en cambio puede suceder que a través de este mecanismo estén obteniendo una nota más baja de la asignatura que la que obtendrían si sólo se calificara con la nota del examen final.

Nota

1. Esta tasa de variación o índice se ha utilizado en contabilidad al tratar de cuantificar las diferencias entre los sistemas contables, o de medir cuantitativamente la armonización en ciertos períodos significativos, principalmente se han usado los denominados índices de medida del conservadurismo, o en grado más amplio, del grado de comparabilidad, e índices de concentración de prácticas contables ([AQ] Sánchez et al., 2002).

Disclosure statement [AQ]

No potential conflict of interest was reported by the authors.

References / Referencias [AQ]

Alcaraz, N., Fernández, M., & Sola, M. (2012). La voz del alumnado en los procesos de evaluación docente universitaria. *Revista Iberoamericana de Evaluación Educativa*, 5(2), 26–39.

Burdett, N., & O'Donnell, S. (2016). Lost in translation? The challenges of educational policy borrowing. *Educational Research*, 58(2), 113–120.

Camacho Miñano, M. M., Urquía Grande, E., Pascual Ezama, D., & Rivero Menéndez, M. J. (2016). Recursos multimedia para el aprendizaje de contabilidad financiera en los grados bilingües. *Educación XXI*, 19(1). doi:10.5944/educXX1.13941 [AQ]

Cano, M. D. (2011). Students' involvement in continuous assessment methodologies: A case study for a distributed information systems course. *IEEE Transactions on Education*, 54(3), 442–451. Art. no. 5590268.

Carless, D. (2007, February). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57–66.

Castejón, F. J., & Santos, M. L. (2011). Percepciones y dificultades en el empleo de metodologías participativas y evaluación formativa en el Grado de Ciencias de la Actividad Física. *Revista Electrónica Interuniversitaria de Formación del Profesorado (REIFOP)*, 14(4). Retrieved from <http://www.aufop.com>

Crisp, B. R. (2007, October). Is it worth the effort? How feedback influences students' subsequent submission of assessable work. *Assessment and Evaluation in Higher Education*, 32(5), 571–581.

Delgado, A. M., & Oliver, R. (2006). La evaluación continua en un nuevo escenario docente. *Revista de Universidad y Sociedad del Conocimiento*, 3(1), 1–13. Retrieved from http://www.uoc.edu/rusc/3/1/dt/esp/delgado_oliver.pdf

- Díaz-Mora, C., García, J. A., & Molina, A. (2016). What is the key to academic success? An analysis of the relationship between time use and student performance. *Cultura y Educación*, 28(1), 157–195.
- Dowell, L. M., Sambell, K., Bazin, V., Penlington, R., Wakelin, D., Wickes, H., & Smailes, J. (2007, May). *Assessment design for learner responsibility*. Retrieved from http://www.reap.ac.uk/reap/reap07/Portals/2/CSL/feast%20of%20case%20studies/Assessment_for_learning_current_practice_exemplars_from_CETL.pdf
- Fraile, A., López, V. M., Castejón, F. J., & Romero, R. (2013). La evaluación formativa en docencia universitaria y el rendimiento académico del alumnado. *Aula Abierta*, 41, 23–34.
- Gámiz Sánchez, V., & Gallego Arrufat, M. J. (2016). Modelo de análisis de metodologías didácticas semipresenciales en educación superior. *Educación XXI*, 19(1), 39–61.
- Gibbs, G. (2006). Why assessment is changing. In C. Bryan, & K. Klegg (Eds.), *Innovative assessment in higher education* (pp.). Abingdon: Routledge.
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. In *Learning and teaching in higher education* (Vol. 1, pp. 3–31). Retrieved from <https://sydney.edu.au/education-portfolio/ei/assessmentresources/pdf/Gibbs%20and%20Simpson.pdf>
- Heitink, M. C., Van Der Kleij, F. M., Veldkamp, B. P., & Schildkamp, K. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review*, 17, 50–62.
- Hernández, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education*, 64, 489–502.
- Inda Caro, M., Álvarez González, S., & Álvarez Rubio, R. (2008). Métodos de evaluación en la enseñanza superior. *Revista de Investigación Educativa*, 26(2), 539–552. Retrieved from <http://www.redalyc.org/articulo.oa?id=283321909014>
- Isaksson, S. (2008). Assess as you go: The effect of continuous assessment on student learning during a short course in archaeology. *Assessment and Evaluation in Higher Education*, 33(1), 1–7.
- Jareño Cebrián, F., & López García, R. (2015). Actividades de evaluación continua-correlación con la calificación de la prueba final y efecto sobre la calificación final. Evidencia en Administración y Dirección de Empresas. *Revista Complutense de Educación*, 26(2), 241–254.

- López-Pastor, V. M. (2011, Enero-Abril). El papel de la evaluación formativa en la evaluación por competencias: Aportaciones de la red de evaluación formativa y compartida en docencia universitaria. *Revista de Docencia Universitaria*, 9(1), 159–173.
- Martínez Clares, P., Pérez Cusó, J., & Martínez Juárez, M. (2016). Las TICs y el entorno virtual para la tutoría universitaria. *Educación XXI*, 19(1), 287–310. [AQ]
- Moro-Egido, A. I. (2016). An assessment of evaluation techniques in higher education: The experience of macroeconomics. *Cultura y Educación*, 28(4), 826–862.
- Palacios Picos, A. (2004). El crédito europeo como motor de cambio en la configuración del Espacio Europeo de la Educación Superior. *Revista Interuniversitaria de Formación del Profesorado*, 18(3), 197–205.
- Pascual Ezama, D., Camacho, M., Ma, M., Urquía Grande, E., & Müller, A. (2011). ¿Son los nuevos criterios de evaluación en el marco del EEES adecuados para valorar el rendimiento académico de los alumnos? Experiencia en Contabilidad Financiera. *Revista de Educación en Contabilidad, Finanzas y Administración de Empresas*, 2, 67–83.
- Poza-Lujan, J. L., Calafate, C. T., Posadas-Yague, J. L., & Cano, J. C. (2016). Assessing the impact of continuous evaluation strategies: Tradeoff between student performance and instructor effort. *IEEE Transactions on Education*, 59(1), 17–23. Art. no. 7086086.
- San Martín Gutiérrez, S., Jiménez Torres, N., & Jerónimo Sánchez-Beato, E. (2016). La evaluación del alumno universitario en el Espacio Europeo de Educación Superior. *Aula Abierta*, 44, 7–14.
- Sánchez-Martín, M. D. P., Pascual-Ezama, D., & Delgado-Jalón, M. L. (2017). Estudiantes mejor informados: Mejores resultados académicos. *Revista de Contabilidad-Spanish Accounting Review*, 20(1), 47–54.
- Trotter, E. (2006, October). Student perceptions of continuous summative assessment. *Assessment and Evaluation in Higher Education*, 31(5), 505–521.
- Vallet-Bellmunt, T., Rivera-Torres, P., & Vallet-Bellmunt, I. (2017). Aprendizaje cooperativo, aprendizaje percibido y rendimiento académico en la enseñanza del marketing. *Educación XXI*, 20(1), 277–297.
- Yorke, M. (2003, June). Formative assessment in higher education: Moves Howard theory and the enhancement of pedagogic practice. *Higher Education*, 45(4), 477–501.

Tabla 1. Características de la muestra.

Número de notas	Curso académico	Convocatoria	Número de pruebas parciales

	2014/15	2015/16	2016/17	1	2	3	2	3	4
Contabilidad Financiera I (2344) ADE 1er cuatrimestre	52	101	96	217	8	24	0	101	148
Contabilidad Financiera II (2351) ADE 2° cuatrimestre	41	179	0	106	92	22	0	124	96
Contabilidad Financiera Superior (2357) ADE 1er cuatrimestre	0	124	49	136	15	22	173	0	0
Introducción a la contabilidad (2210) Ciencias Políticas 1er cuatrimestre	63	0	0	31	29	3	63	0	0

Fuente: Elaboración propia.

Tabla 2. Tabla de contingencia presentados/no presentados al examen final en base a haber superado o no la evaluación continua.

Aprueba evaluación continua	Primera convocatoria Presenta examen, n (%)	
	Sí	No
No	397 (56.3)	833 (86.3)
Sí	308 (43.7)	132 (13.7)

Nota: $\chi^2(1) = 189.053, p < .001$

Fuente: Elaboración propia.

Tabla 3. Estadísticos descriptivos de los índices.

	N	Mínimo	Máximo	Media	Desv. Típ.
Índice global	705	-56	96.1	4.1	25.17
Índice positivo	358	0	96.1	10.5	19.43
Índice negativo	347	-56	0	-13.2	12.40

Fuente: Elaboración propia.

Tabla 4. ANOVA del índice y número de pruebas parciales.

Número de pruebas	N	Media	Desviación estándar	Error estándar	95% del intervalo de confianza de la media		Mínimo	Máximo
					Límite inferior	Límite superior		
2	236	1.7148a	23.70139	1.54283	4.7544	1.3247	55.90	96.10
3	225	10.1973b	31.48266	2.09884	6.0613	14.3333	42.90	84.60
4	244	3.9840c	17.64841	1.12982	1.7585	6.2095	42.90	81.90
Total	705	4.0593	24.17171	0.94802	2.1980	5.9206	55.90	96.10

Nota: ($F(2) = 13.352, p < .001$); a-b-c. Comparaciones dos a dos post hoc (Tukey).

Diferentes letras indican diferencias estadísticamente significativas a nivel $p < .05$.

Fuente: Elaboración propia.

Tabla 5. ANOVA del índice por curso académico.

Curso académico	N	Media	Desviación estándar	Error estándar	95% del intervalo de confianza de la media		Mínimo	Máximo
					Límite inferior	Límite superior		
2014–15	156	14.7205a	40.58955	3.24977	8.3010	21.1401	55.90	84.60
2015–16	404	1.8040b	15.99224	0.79564	3.3681	0.2398	42.90	96.10
2016–17	145	8.9255a	19.10849	1.58687	5.7889	12.0621	42.90	81.90
Total	705	4.0593	25.17171	0.94802	2.1980	5.9206	55.90	96.10

Nota: ($F(2) = 29.935, p < .001$); a-b-c. Comparaciones dos a dos post hoc (Tukey).

Diferentes letras indican diferencias estadísticamente significativas a nivel $p < .05$.

Fuente: Elaboración propia.

Tabla 6. ANOVA del índice por asignatura.

Asignatura	N	Media	Desviación estándar	Error estándar	95% del intervalo de confianza de la media		Mínimo	Máximo
					Límite inferior	Límite superior		
2210 Introducción Contabilidad	63	11.4810a	28.83874	3.63334	18.7439	4.2180	55.90	61.70
2344 Contabilidad Financiera I	249	3.8293b,c	18.00591	1.14108	1.5819	6.0768	42.90	81.90
2351 Contabilidad Financiera II	220	10.5136c	31.44912	2.12030	6.3348	14.6924	42.90	84.60
2357 Contabilidad	173	1.8416b	20.49561	1.55825	1.2341	4.9174	42.90	96.10

Asignatura	N	Media	Desviación estándar	Error estándar	95% del intervalo de confianza de la media		Mínimo	Máximo
					Límite inferior	Límite superior		
Financiera Superior								
Total	705	4.0593	25.17171	0.94802	2.1980	5.9206	55.90	96.10

Nota: ($F(3) = 14.017, p < .001$); a-b-c. Comparaciones dos a dos post hoc (Tukey).

Diferentes letras indican diferencias estadísticamente significativas a nivel $p < .05$.

Fuente: Elaboración propia.

Tabla 7. ANOVA del índice por convocatoria.

Convocatoria	N	Media	Desviación estándar	Error estándar	95% del intervalo de confianza de la media		Mínimo	Máximo
					Límite inferior	Límite superior		
1	490	7.4092a	26.90854	1.21560	5.0207	9.7976	55.90	96.10
2	144	3.0104b	18.62238	1.55186	6.0780	0.0571	42.90	62.90
3	71	4.7211b	18.60650	2.20818	9.1252	0.3170	42.90	60.50
Total	705	4.0593	25.17171	0.94802	2.1980	5.9206	55.90	96.10

Nota: ($F(2) = 14.904, p < .001$); a-b-c. Comparaciones dos a dos post hoc (Tukey).

Diferentes letras indican diferencias estadísticamente significativas a nivel $p < .05$.

Fuente: Elaboración propia.

Tabla 8. Estimación de los modelos de regresión.

	Variable dependiente					
	Nota_examen_final		Indice_global		Nota_eva_cont	
	Beta(SE)	p	Beta(SE)	p	Beta(SE)	p
N = 705						
Constante	2.596(0.847)	.002	78.506(9.023)	.000	0.111(0.882)	.900
Nota_eva_cont	0.555(0.038)	.000	-	-	-	-
Número_pruebas	0.858(0.233)	.000	14.677(2.429)	.000	1.291(0.237)	.000
Curso15_16	2.445(0.250)	.000	39.228(2.700)	.000	0.609(0.259)	.019
Curso16_17	0.794(0.273)	.004	15.770(2.962)	.000	0.175(0.284)	.538
Asign2351	0.394(0.211)	.062	13.931(2.259)	.000	0.520(0.219)	.018

	Variable dependiente					
	<i>Nota_examen_final</i>		<i>Indice_global</i>		<i>Nota_eva_cont</i>	
<i>N = 705</i>	Beta(SE)	<i>p</i>	Beta(SE)	<i>p</i>	Beta(SE)	<i>p</i>
<i>Asign2357</i>	1.106(0.402)	.006	14.860(4.244)	.000	2.153(0.410)	.000
<i>Asign2210</i>	3.230(0.544)	.000	61.079(5.820)	.000	1.925(0.561)	.001
<i>Conv2</i>	-0.017(0.242)	.943	0.927(2.620)	.724	-0.381(0.252)	.131
<i>Conv3</i>	-0.288(0.269)	.286	0.416(2.850)	.884	-0.991(0.278)	.000
<i>R² ajustado (%)</i>	38.1		32		6.6	
Modelo	$F(9) = 43.9^{***}$		$F(8) = 42.391^{***}$		$F(8) = 6.777^{***}$	

Nota: ****p*-valor < .01; ***p*-valor < .05; **p*-valor < .10. Los errores estándar aparecen entre paréntesis.

Fuente: Elaboración propia.

Tabla 9. Correlaciones de muestras emparejadas.

		Correlación
<i>Par 1</i>	<i>Nota_eva_cont & Nota_exa_fin</i>	.497***
<i>Par 2</i>	<i>Nota_eva_cont & Nota_fin_asig</i>	.551***
<i>Par 3</i>	<i>Nota_exa_fin & Nota_fin_asig</i>	.913***
*** <i>p</i> -valor < .01	** <i>p</i> -valor < .05	* <i>p</i> -valor < .1

Fuente: Elaboración propia.

Tabla 10. Prueba de muestras emparejadas.

		Diferencias emparejadas					<i>t</i>	<i>gl</i>	Sig. (bilateral)
		Media	Desviación típica	Error	95% intervalo confianza				
					Inferior	Superior			
<i>Par 1</i>	<i>Nota_eva_cont</i>	0.27254	2.15830	0.08459	0.10643	0.43864	3.222	650	.001

		Diferencias emparejadas							
					95% intervalo confianza				
		Media	Desviación típica	Error	Inferior	Superior	<i>t</i>	<i>gl</i>	Sig. (bilate ral)
	& <i>Nota_exa_fin</i>								
<i>Par 2</i>	<i>Nota_eva_cont</i> & <i>Nota_fin_asig</i>	1.70024	2.35813	0.06471	1.57329	1.82718	26.275	1,327	.000
<i>Par 3</i>	<i>Nota_exa_fin</i> & <i>Nota_fin_asig</i>	0.06527	0.94719	0.03567	0.13531	0.00476	1.83	704	.068
*** <i>p</i> -valor<.01	** <i>p</i> -valor<.05	* <i>p</i> -valor<.1							

Fuente: Elaboración propia.

Figura 1. Media del índice global por número de pruebas.

Fuente: Elaboración propia.

Figura 2. Media del índice global por curso académico.

Fuente: Elaboración propia.

Figura 3. Media del índice global por asignatura.

Fuente: Elaboración propia.

Figura 4. Media del índice global por convocatoria.

Fuente: Elaboración propia.

Figure 1. Mean of the global index by number of evaluations.

Source: Authors.

Mean of the global index

Mean

Number of evaluations

Figure 2. Mean of the global index by academic year.

Source: Authors.

Mean of the global index

Academic year

Figure 3. Mean of the global index by course.

Source: Authors.

Mean of the global index

Course

Figure 4. Mean of the global index by exam period.

Source: Authors.

Mean of the global index

Mean

Exam period

Table 1. Characteristics of the sample.

Number of marks	Academic year			Exam period			Number of partial evaluations		
	2014/15	2015/16	2016/17	1	2	3	2	3	4
Financial Accounting I (2344) BBA 1st semester	52	101	96	217	8	24	0	101	148
Financial Accounting II (2351) BBA 2nd semester	41	179	0	106	92	22	0	124	96
Upper-Level Financial Accounting (2357) BBA 1st semester	0	124	49	136	15	22	173	0	0
Introduction to Accounting (2210) Political Science 1st semester	63	0	0	31	29	3	63	0	0

Source: Authors.

Table 2. Contingency table of students who sat/did not sit for the final exam based on whether or not they had passed the continuous assessment.

Continuous assessment evaluations	First exam period Sat for exam, <i>n</i> (%)	
	Yes	No
No	397 (56.3)	833 (86.3)
Yes	308 (43.7)	132 (13.7)

$$X^2(1) = 189.053, p < .001$$

Source: Authors.

Table 3. Descriptive statistics of the indexes.

	<i>N</i>	Minimum	Maximum	Mean	St. Dev.

	<i>N</i>	Minimum	Maximum	Mean	St. Dev.
Global index	705	56	96.1	4.1	25.17
Positive index	358	0	96.1	10.5	19.43
Negative index	347	56	0	13.2	12.40

Source: Authors.

Table 4. ANOVA of the index and number of partial evaluations.

					95% confidence interval of the mean			
Number of evaluations	<i>N</i>	Mean	Standard deviation	Standard error	Lower limit	Upper limit	Minimum	Maximum
2	236	1.7148a	23.70139	1.54283	4.7544	1.3247	55.90	96.10
3	225	10.1973b	31.48266	2.09884	6.0613	14.3333	42.90	84.60
4	244	3.9840c	17.64841	1.12982	1.7585	6.2095	42.90	81.90
Total	705	4.0593	24.17171	0.94802	2.1980	5.9206	55.90	96.10

Note: ($F(2) = 13.352, p < .001$); a-b-c. Two-by-two post hoc comparisons (Tukey).

Different letters indicate statistically significant differences at a level of $p < .05$.

Source: Authors.

Table 5. ANOVA of the index by academic year.

					95% confidence interval of the mean			
Academic year	<i>N</i>	Mean	Standard deviation	Standard error	Lower limit	Upper limit	Minimum	Maximum
2014–15	156	14.7205a	40.58955	3.24977	8.3010	21.1401	55.90	84.60
2015–16	404	1.8040b	15.99224	0.79564	3.3681	-0.2398	42.90	96.10
2016–17	145	8.9255a	19.10849	1.58687	5.7889	12.0621	42.90	81.90
Total	705	4.0593	25.17171	0.94802	2.1980	5.9206	55.90	96.10

Note: ($F(2) = 29.935, p < .001$); a-b-c. Two-by-two post hoc comparisons (Tukey).

Different letters indicate statistically significant differences at a level of $p < .05$.

Source: Authors.

Table 6. ANOVA of the index by course.

					95% confidence interval of the mean			
Course	<i>N</i>	Mean	Standard deviation	Standard error	Lower limit	Upper limit	Minimum	Maximum
2210 Introduction to	63	11.4810a	28.83874	3.63334	18.7439	4.2180	55.90	61.70

Course	N	Mean	Standard deviation	Standard error	95% confidence interval of the mean		Minimum	Maximum
					Lower limit	Upper limit		
Accounting								
2344 Financial Accounting I	249	3.8293b,c	18.00591	1.14108	1.5819	6.0768	-42.90	81.90
2351 Financial Accounting II	220	10.5136c	31.44912	2.12030	6.3348	14.6924	-42.90	84.60
2357 Upper-level Financial Accounting	173	1.8416b	20.49561	1.55825	-1.2341	4.9174	-42.90	96.10
Total	705	4.0593	25.17171	0.94802	2.1980	5.9206	-55.90	96.10

Note: ($F(3) = 14.017, p < .001$); a-b-c. Two-by-two post hoc comparisons (Tukey).

Different letters indicate statistically significant differences at a level of $p < .05$.

Source: Authors.

Table 7. ANOVA of the index by exam period.

Exam period	N	Mean	Standard deviation	Standard error	95% confidence interval of the mean		Minimum	Maximum
					Lower limit	Upper limit		
1	490	7.4092a	26.90854	1.21560	5.0207	9.7976	-55.90	96.10
2	144	-3.0104b	18.62238	1.55186	-6.0780	0.0571	-42.90	62.90
3	71	-4.7211b	18.60650	2.20818	-9.1252	-0.3170	-42.90	60.50
Total	705	4.0593	25.17171	0.94802	2.1980	5.9206	-55.90	96.10

Note: ($F(2) = 14.904, p < .001$); a-b-c. Two-by-two post hoc comparisons (Tukey).

Different letters indicate statistically significant differences at a level of $p < .05$.

Source: Authors.

Table 8. Estimation of the regression analyses.

	Dependent variable					
	<i>Exam_mark_final</i>		<i>Overall_index</i>		<i>CA_mark</i>	
<i>N = 705</i>	Beta(SE)	<i>p</i>	Beta(SE)	<i>p</i>	Beta(SE)	<i>p</i>
<i>Constant</i>	-2.596(0.847)	.002	78.506(9.023)	.000	-0.111(0.882)	.900
<i>CA_mark</i>	0.555(0.038)	.000	-	-	-	-

	Dependent variable					
	<i>Exam_mark_final</i>		<i>Overall_index</i>		<i>CA_mark</i>	
<i>N = 705</i>	Beta(SE)	<i>p</i>	Beta(SE)	<i>p</i>	Beta(SE)	<i>p</i>
<i>Number_evals</i>	0.858(0.233)	.000	-14.677(2.429)	.000	1.291(0.237)	.000
<i>AY15_16</i>	2.445(0.250)	.000	-39.228(2.700)	.000	0.609(0.259)	.019
<i>AY16_17</i>	0.794(0.273)	.004	-15.770(2.962)	.000	0.175(0.284)	.538
<i>Class2351</i>	-0.394(0.211)	.062	13.931(2.259)	.000	0.520(0.219)	.018
<i>Class2357</i>	1.106(0.402)	.006	-14.860(4.244)	.000	2.153(0.410)	.000
<i>Class2210</i>	3.230(0.544)	.000	-61.079(5.820)	.000	1.925(0.561)	.001
<i>Ex_Per2</i>	-0.017(0.242)	.943	0.927(2.620)	.724	-0.381(0.252)	.131
<i>Ex_Per3</i>	-0.288(0.269)	.286	0.416(2.850)	.884	-0.991(0.278)	.000
Adjusted R² (%)	38.1		32		6.6	
Model	<i>F</i> (9) = 43.9***		<i>F</i> (8) = 42.391***		<i>F</i> (8) = 6.777***	

Note: ****p*-value < .01; ***p*-value < .05; **p*-value < .10. Standard errors appear between parentheses.

Source: Authors.

Table 9. Correlations of paired samples.

		Correlation
<i>Pair 1</i>	<i>CA_mark & Final_ex_mark</i>	.497***
<i>Pair 2</i>	<i>CA_mark & Class_mark</i>	.551***
<i>Pair 3</i>	<i>Final_ex_mark & Class_mark</i>	.913***
*** <i>p</i> -value < .01	** <i>p</i> -value < .05	* <i>p</i> -value < .1

Source: Authors.

Table 10. Evaluation of paired samples.

		Paired differences			

					95% confidence interval				
		Mean	Standard deviation	Error	Higher	Lower	<i>t</i>	<i>df</i>	Sig. (bilateral)
<i>Pair 1</i>	<i>CA_mark & Final_ex_mark</i>	0.27254	2.15830	0.08459	0.10643	0.43864	3.222	650	.001
<i>Pair 2</i>	<i>CA_mark & Class_mark</i>	1.70024	2.35813	0.06471	1.57329	1.82718	26.275	1,327	.000
<i>Pair 3</i>	<i>Final_ex_mark & Class_mark</i>	-0.06527	0.94719	0.03567	-0.13531	0.00476	-1.83	704	.068
*** <i>p</i> -value<.01	** <i>p</i> -value<.05	* <i>p</i> -value<.1							

Source: Authors.