

Effect Of Continuous Assessment Tests On Overall Student Performance In The Case Of The Spanish National Distance Education University (UNED)

María Dolores Reina-Paz, Universidad Nacional de Educación a Distancia (UNED), Spain
Ainhoa Rodríguez-Oromendía, Universidad Nacional de Educación a Distancia (UNED), Spain
Claudia Sevilla-Sevilla, Universidad Nacional de Educación a Distancia (UNED), Spain

ABSTRACT

This study aims to analyze the effect of the use of different types of continuous assessment tests on students' participation, performance and final grades. To this end, it examines the case of a second-year course from the degree program in Business Administration and Management offered by the Faculty of Economics at the Universidad Nacional de Educación a Distancia (National Distance Education University, UNED) in Spain during the first two academic years it was available (2011/12 and 2012/13). This period offered two distinct scenarios, with a total of 1,450 students participating the first year and 1,809 in the second, due to the use of two online continuous assessment tests with different scoring systems. The goal was to analyze the influence of each of the continuous assessment systems used on student participation in both the continuous assessment test itself and the final exam, as well as on the grades earned on these tests, and student's overall performance in the course. As a result of this research, conclusions were reached regarding the educational nature of these tests and their suitability as a true tool for monitoring and assessment at a distance university.

Keywords: Continuous Assessment Tests; Distance Education; Performance; Participation; Learning

1. INTRODUCTION

The new teaching framework established as a result of the European Higher Education Area (EHEA) adaptation process requires the reformulation of teaching methods, which must now be based not only on teaching, but also learning.¹ This model places students at the center of the learning process, and the European credit thus becomes a unit for assessing the total volume of a student's work. Moreover, assessments become the indisputable keystone of the system, since students earn credits by passing the assessment system used for each course (Delgado, 2005).

This process of convergence toward the EHEA entails certain methodological changes whose practical application can be challenging for distance universities and whose implications for teaching quality must thus be studied and reviewed. In this regard, learning assessments are one of the issues attracting most interest in the field of teaching and one of the most studied and addressed by teachers and experts alike.²

It is possible to identify three distinct types of assessments, depending on when and why they are carried out: diagnostic, or *initial*, assessments; formative, or *continuous*, assessments; and summative, or *final*, assessments.

¹ See, for example, Margalef et al. (2007) with regard to innovative experiences in teaching at the university level.

² On the different assessment processes, assessment challenges, and learning to assess, see, among others: Cabra Torres (2008), Ávila Cañadas et al. (2010), García González (2004), Barberà Gregori & de Martín Rojo (2009), and Vargas Quintero (2008).

Diagnostic, or *initial*, assessments are intended to bring the starting point for training into line with students' prior knowledge and skills. This type of assessment is unnecessary in the case of courses forming a part of a given qualification or degree program, as such courses are scheduled within a curriculum that takes into account the skills and knowledge that students are expected to acquire each year. Formative, or *continuous*, assessments are used to provide feedback to improve both student learning and teaching quality. These types of assessments must be carried out during the learning process. The last type of assessment, the summative, or *final*, assessment, is administered at the end of the learning process in order to grade the student. This type of assessment is purely quantitative in nature, as it has no bearing on the learning process itself, but rather only on the certification thereof or the awarding of a final grade.

One of the most challenging aspects for universities that exclusively use distance-learning methods is the *continuous assessment*,³ as in most cases there is no direct contact between students and professors, making the monitoring required for continuous assessment processes much more demanding and time-consuming. Especially in cases where there is a large number of students (as in the case of this study), a continuous assessment system increases the workload and time required of teachers, due to the need not only to prepare the different planned activities, but also to correct and grade them (Delgado, 2006). One way to make the continuous assessment system effective is to combine it with self-assessment techniques in which students participate in the decision of how to assess their own work. Students thus internalize the assessment criteria, find alternative solutions to their papers, become more involved in the learning process, and acquire the habit of critically thinking about their own work. A self-assessment system that is supported in e-learning environments and designed not to overburden students but rather to be understood as part of the learning process enables fulfillment of the goals of continuous assessment without creating an excessive workload for teachers (Mosqueira-Rey, 2010).

In the case of the course studied here, the continuous assessment process was carried out using two types of tools: self-assessment tests offered to students throughout the course in the virtual classroom; and an online continuous assessment test (CAT) offered to students at a specific time and date, prior to the start of the exam period. The purpose of this test was to obtain a score for the student's continuous assessment to date, for use not only by the teacher but also by the student, as the test would also serve as both a means of self-assessment (the computer system reported the grade earned immediately upon completion of the test) and a trial run for the course's final exam.

The study described here focused on this CAT with a view to offering an initial analysis of the effect of such tests on student performance and on the results of the course in terms of the participation rate and grades. This type of CAT was used the first two years in which the course – *Introduction to Marketing*, a first-semester course in the second year of the Business Administration and Management degree program – was offered. The test design and conditions under which the test was offered were the same in both years. Specifically, the tests were optional and consisted of 10 multiple-choice questions on all the material covered. Students could take them online using any computer through the virtual classroom hosted on the UNED website. The test was enabled for 12 hours beginning at a specific time and date. Upon accessing the test, students had a maximum of 30 minutes to complete it. As soon as the test had been completed, the computer system gave the students their score. This score was calculated as follows: students earned 0.1 points for each correct answer, and no points for wrong answers and questions left blank. Students could thus obtain a maximum score of 1 point.

However, there was one difference between the tests: different assessment systems were used to determine their weight in the student's overall grade (*final grade*). Both years, students were required to pass an in-person multiple-choice exam with a score of at least 5 for the CAT to be included in the calculation of their final grade for the course. Once this condition had been met, the CAT counted for 1 point (to be added to the final grade) in the 2011/12 school year, provided the student scored at least 0.5 points on it. In contrast, in the 2012/13 school year, only the actual points that students had earned on the CAT were added to their grade on the final exam to calculate the final grade for the course. The two assessment systems were thus quite different and offered different incentives to students. The system used the first year was primarily designed to encourage students to participate in the CAT

³ On different aspects of continuous assessment and studies at different universities, see: Delgado García (2005), Delgado & Oliver (2006), López (2007), Montoro et al. (2011), Mosqueira-Rey (2010), and Ruiz Conde (2010).

and gave relatively little importance to their actual score (provided they reached the minimum score of 0.5 points); in contrast, the system used the second year incentivized students both to participate in the CAT and to earn their best possible score on it.

This paper aims to serve as an initial step in the study of the effects that CATs can have on student performance during the learning process and on their final grades for a course.

2. RESEARCH METHOD AND DESCRIPTIVE ANALYSIS OF THE SAMPLE

In this initial stage of the study, the most relevant data available in the database for the UNED’s educational platform Alf was examined for both of the academic years under study. Three variables were established with regard to the online CAT: an initial variable called “*CAT Status*,” covering the three possible degrees of completion of each student’s CAT (Taken, Not taken, Incomplete); a second variable called “*CAT System*,” indicating the assessment system used for completed CATs each year; and a third variable called “*CAT Time*,” indicating the time each student took to complete the CAT in minutes. Three variables were also established with regard to the *results*, covering the grades students earned on the different assessment processes (the CAT, the multiple-choice exam, and the final assessment for the course). Table 1 shows the variables studied, along with their operational definitions.

Table 1: Variables Used in the Analysis

CAT Status	Taken	The student took the CAT
	Not taken	The student did not take the CAT
	Incomplete	The student began the CAT but did not finish it
CAT System	System used to determine the value to be added for the CAT to the final grade for the course. In 2012: CAT value = 1 point (subject to a CAT score ≥ 5) In 2013: CAT value = CAT score/10	
CAT Time	Minutes taken by the student to complete the CAT.	
CAT Grade	Grade earned on the CAT.	
Exam Grade	Grade earned on the in-person assessment test. This test was a multiple-choice final exam for the course.	
Final Grade	Final grade earned (EXAM GRADE + CAT value) The CAT value was only included for students who scored at least 5 on the EXAM. ⁴	

The sample consisted of 3,259 participants, with somewhat more than half (55.5%) belonging to the 2012/13 school year, and the rest (44.5%) belonging to the 2011/12 school year. The sample had a fairly equal gender distribution, consisting of nearly 54% men and 46% women. With regard to age, while the sample included several far-out outliers, they accounted for only a negligible percentage. They could thus be excluded with virtually no effect on the descriptive analyses of this variable, which showed a strong positive skewness implying a significant deviation from the Gaussian model. The mean age was 33.17 years old, with a standard deviation of 8.02 years within an age range of 19 to 77. As for the size of the Associated Centers to which students belonged, 57.3% of the students enrolled in the course were registered at large centers (with 30 or more enrolled students), 27.8% at medium-sized centers (between 10 and 29 students), and the rest of the sample (14.8%) at small centers (less than 10 students). Table 2 offers a descriptive summary of the study sample.

Table 2: Descriptive Analysis of the Sample (N = 3,259)

Descriptor			
Students Enrolled	2012: 44.5% (1,450 students)		2013: 55.5% (1,809 students)
Sex	Men: 53.8% (1,752 students)		Women: 46.2% (1,507 students)
Age	33.17 years old ± 8.02 (C.I.: 32.90-33.45) Range: 19-77		
Center	Large: 57.3% (1,869 students)	Medium: 28.8% (907 students)	Small: 14.8% (483 students)

⁴ Scores of 4.8 are automatically rounded up to 5. In these cases, the value of the CAT is not included in the calculation of the final score (FINAL GRADE).

With regard to the CAT status, only 6 registered students (0.2%) began the test without completing it. Given the negligible nature of this number, these students were excluded from the rest of the analysis. The final sample thus consisted of 3,253 participants. Of this group, 65.8% of the total sample for both academic years did not take the CAT, such that the overall participation rate for both years was 34.2%. An analysis of participation by year shows an increase of 15.28% between 2012 (when participation stood at 31.4%) and 2013 (when it stood at 36.2%). As for the time taken to complete the CAT, the distribution of the participants was relatively homogeneous: times ranged from 1 to 23 minutes, such that the variable differs significantly from the normal distribution. There were no marked outliers, and only 1.3% of the participants who took the CAT took longer than the aforementioned 23 minutes. The average time spent was 11.70 minutes, with a standard deviation of 6.47 minutes. A comparison by year (Table 3) showed no significant differences.

Table 3: Description of the Sample Analyzed Based on the CAT (N = 3,253)

Variable	Descriptor			Significance
CAT Status	Taken: 34.0% (1,107 students)	Not taken: 65.8% (2,146 students)	Incomplete: 0.2% (6 students)	
2012	Taken: 31.4% (454 students)	Not taken: 68.6% (993 students)	-	0.004**
2013	Taken: 36.2% (653 students)	Not taken: 63.8% (1,153 students)		
CAT Time	11.70 minutes \pm 6.47 (C.I.: 11.32-12.08) Range: 1-23			
2012	12.07 \pm 6.46			0.106 ^{NS}
2013	11.44 \pm 6.47			

NS = Not significant ($p > 0.050$), * Significant at 5%, ** Highly significant at 1%

3. PROPOSED HYPOTHESES

In order to respond to the initial questions raised in this study, the following initial hypotheses were proposed:

H1: Student participation in the CAT influences student participation in the in-person multiple-choice exam.

To test this hypothesis, a dichotomous variable was generated called “SAT EXAM.” Some 63% of all participants (2,048 students) sat the in-person multiple-choice exam. Of this group, 92.6% had previously participated in the CAT. Table 4 shows the high significance of this variable, indicating the existence of a relationship between students’ participation in each test (the CAT and the multiple-choice exam). This hypothesis can thus be accepted.

Table 4: Comparison of Percentage of Students Sitting the Exam by CAT STATUS and CAT SYSTEM

Variable		Percentage of Students Sitting The Exam		χ^2	GI	Sig.
SAT EXAM		Yes: 63% (2,048)	No: 37% (1,205)			
CAT STATUS	Taken (N = 1,107)	Yes: 92.6% (1,025)	No: 7.4% (82)	631,935	1	P = 0.000**
	Not taken (N = 2,146)	Yes: 47.7% (1,023)	No: 52.3% (1,123)			
CAT SYSTEM	2012			327,351	1	P = 0.000**
	Took CAT (N = 454)	Yes: 95.1% (432)	No: 4.9% (22)			
	Did not take CAT (N = 993)	Yes: 45.1% (448)	No: 54.9% (545)			
	2013			305,865	1	P = 0.000**
Took CAT (N = 653)	Yes: 90.8% (593)	No: 9.2% (60)				
	Did not take CAT (N = 1,153)	Yes: 49.9% (575)	No: 50.1% (578)			

NS = Not significant ($p > 0.050$), * Significant at 5%, ** Highly significant at 1%

This effect on student participation in both tests was found in both periods (2011/12 and 2012/13), although a certain difference can be seen between one year and the next, as the number of students sitting the exam after participating in the CAT in 2013 was down 5% from 2012. One might speculate that this fact could be due to the existence of a significant difference between the grades students earned on the CAT in accordance with the assessment system used. Were such a difference to exist in the grades depending on the CAT assessment system

used, students' decision to sit the exam might have been influenced by their grade on the CAT. This possibility will be examined in the discussion of hypothesis *H3* below.

H2: Students who take the CAT earn higher final grades for the course.

A total of 2,054 students, some 63% of those in the database, sat the in-person exam. The mean grade on the in-person exam (EXAM GRADE) was 6.12, with sufficient negative skewness (toward higher grades) to entail a significant difference with the normal bell curve. The mean grade of the 1,025 participants who had taken the CAT was 6.67, higher than the mean grade of the 1,023 students who had not, which was 5.58. This difference proved to be significant (Table 6) and makes it possible to estimate a difference of nearly 1.1 points in the grade on the in-person exam, with students who have taken the CAT scoring higher.

The mean final grade was 6.50 and again showed sufficient negative skewness (toward higher grades) to differ significantly from the Gaussian bell curve. As shown in Table 6, the mean grade of the 1,025 students who had taken the CAT was 7.42, higher than the mean grade of the 1,023 students who had not (5.58). This difference is statistically significant and makes it possible to estimate a difference of nearly 2 points (1.84) in the final grade, once again with students who had taken the CAT scoring higher. As a result of all of the above, hypothesis *H2* can be accepted.

Table 5: Mean Difference Test: Performance by CAT STATUS

Variable	Mean ± S.D. (C.I. 95%)	T	GI	Mean Dif.	Sig.
Exam Grade (N = 2,054)	6.12 ± 2.17 (C.I.: 6.03-6.22)				
Took CAT (N = 1025)	6.67 ± 1.96 (C.I.: 6.55-6.79)	11.75	2,046	1.09	P = 0.000**
Did not take CAT (N = 1023)	5.58 ± 2.24 (C.I.: 5.44-5.71)				
Final Grade (N = 2,054)	6.50 ± 2.41 (C.I.: 6.39-6.60)				
Took CAT (N = 1025)	7.42 ± 2.22 (C.I.: 7.28-7.56)	18.72	2,046	1.84	P = 0.000**
Did not take CAT (N = 1023)	5.58 ± 2.24 (C.I.: 5.44-5.72)				

NS = Not significant ($p > 0.050$), * Significant at 5%, ** Highly significant at 1%

A second step was carried out (using the 2-way fixed effects multivariate ANOVA method) to study the joint effect of participation in the CAT and the assessment system used each year on student performance. The results showed significant effects on student performance for each variable taken independently (as shown in Tables 5 and 6), but not when the variables were taken together ($p = 0.094$).

H3: The assessment system used for the CAT influences the results obtained by students.

This hypothesis was proposed for the purpose of verifying whether the different assessment systems used for the CAT bear any relation to students' performance (the grades earned) on the different tests (CAT, multiple-choice exam and final grade).

Table 6 summarizes the results obtained for this new hypothesis. As can be seen, this factor was found to be highly significant in relation to each of the three variables related to students' results (grade on the CAT, grade on the in-person multiple-choice exam, and final grade for the course). It can thus be concluded that the assessment system used for the CAT influences not only the grade earned on the CAT itself, but also student performance on the in-person exam and, obviously, the final grade for the course. An analysis of the mean grades earned on the CAT shows that the mean grade⁵ in 2012 (8.99) was lower than the mean grade in 2013 (9.13). This was most likely due to the fact that under the assessment system used for the CAT in 2012, students only needed to score a 5 on the test to receive 1 full point toward the calculation of their final grade for the course. They thus had no incentive to earn a score higher than 5 on the CAT; hence, the lower mean grade that year. The opposite was true for the mean grade on the in-person multiple-choice exam. In that case, the mean grade was higher in 2012, most likely due to the greater incentive for students resulting from the fact that a much more favorable assessment system was used for the CAT than in 2013.

⁵ For the CAT, the grade was calculated awarding 1 point for each question the student answered correctly out of the total of 10 questions included on the test.

Table 6: Mean Difference Test: Performance by CAT SYSTEM (2012 and 2013)

Variable	Mean \pm S.D. (C.I. 95%)	T	GI	Mean Dif.	sig.
CAT Grade (N = 1,107)	9.10 \pm 1.44 (C.I.: 9.02-9.19)				
2012 System (N = 454)	8.99 \pm 1.52 (C.I.: 8.85-9.13)	2.19	1,105	0.19	P = 0.014 *
2013 System (N = 653)	9.18 \pm 1.38 (C.I.: 9.08-9.29)				
Exam Grade (N = 2,054)	6.12 \pm 2.17 (C.I.: 6.03-6.22)				
2012 System (N = 883)	6.44 \pm 2.12 (C.I.: 6.30-6.58)	5.83	2,052	0.56	P = 0.000**
2013 System (N = 1,171)	5.88 \pm 2.18 (C.I.: 5.76-6.00)				
Final Grade (N = 2,054)	6.50 \pm 2.41 (C.I.: 6.39-6.60)				
2012 System (N = 883)	6.83 \pm 2.33 (C.I.: 6.67-6.98)	5.44	2,052	0.58	P = 0.000**
2013 System (N = 1,171)	6.25 \pm 2.44 (C.I.: 6.10-6.39)				

NS = Not significant ($p > 0.050$), * Significant at 5%, ** Highly significant at 1%

Given the similarity between the values obtained for the exam grade and final grade variables, the relationship between these variables was analyzed. Unsurprisingly, this analysis yielded a nearly perfect linear regression value (0.985; significance $p < 0.000$), indicating that 97% of all variation in the final grade is associated with the grade on the in-person multiple-choice exam, leaving little room for the effect of other factors. Accordingly, performance can be studied using either of the two grades.

H4: The time taken to complete the CAT influences performance on the test.

Pearson's correlation coefficient was determined. In this case, it was close to zero (-0.044) and, thus, logically not significant with $p > 0.05$ ($P = 0.145$). This hypothesis was thus rejected as there was no relationship between the time taken by students to complete the CAT and the grade earned. This might lead one to think that the CAT had a low level of difficulty, as students obtained similar grades, most of them quite high, regardless of how much time they spent taking the test. If that were true, most students would spend little time on it, basically the minimum amount considered necessary to read the 10 questions making it up carefully, that is, 3 minutes. As shown in Table 7, out of the 1,107 students who took the test, only 126, or 11.38%, spent 3 minutes or less answering the questions. If the reason this group of students was "faster" than average (11.70 minutes) were the low difficulty level, for them, of the CAT, their grades on it could be expected to be much higher than those of the other 981 students, who needed more time to complete it. Hypothesis *H5* was proposed to test this possibility.

H5: Students who complete the CAT in 3 minutes or less earn significantly higher grades than students who take longer.

To test this hypothesis, a dichotomous variable was created based on the time taken to complete the CAT (CAT TIME \leq 3).

Table 7: Mean Difference Test: Performance by CAT TIME \leq 3 Minutes

Variable	Mean \pm S.D. (C.I. 95%)	T	GI	Mean Dif.	Sig.
CAT Grade	9.10 \pm 1.44 (C.I.: 9.02-9.19)				
CAT Time \leq 3 min. (n = 126)	9.18 \pm 1.37 (C.I.: 8.94-9.43)	0.66	1,105	0.09	P = 0.512 ^{NS}
CAT Time $>$ 3 min. (n = 981)	9.09 \pm 1.45 (C.I.: 9.00-9.19)				
Exam Grade	6.12 \pm 2.17 (C.I.: 6.03-6.22)				
CAT Time \leq 3 min. (n = 115)	6.47 \pm 2.20 (C.I.: 6.06-6.87)	1.18	1,023	0.22	P = 0.238 ^{NS}
CAT Time $>$ 3 min. (n = 910)	6.69 \pm 1.93 (C.I.: 6.57-6.82)				
Final Grade	6.50 \pm 2.41 (C.I.: 6.39-6.60)				
CAT Time \leq 3 min. (n = 115)	7.17 \pm 2.50 (C.I.: 6.71-7.37)	1.26	1,023	0.28	P = 0.207 ^{NS}
CAT Time $>$ 3 min. (n = 910)	7.45 \pm 2.18 (C.I.: 7.31-7.59)				

NS = Not significant ($p > 0.050$), * Significant at 5%, ** Highly significant at 1%

An analysis of the mean differences between the subjects of each group with regard to both the CAT and the in-person exam and final grade (Table 7) revealed no statistically significant differences with $p > 0.05$. Hypothesis *H5* can thus be rejected. This in turn might suggest that these "fast" students managed to finish the CAT in record time not due to their high level of mastery of the subject matter (as indicated by both their mean scores on

the multiple-choice exam and their final grades for the course), but rather because they had obtained information about the answers to the questions on the CAT from classmates who had already taken it.

Returning to the result of the test of hypothesis H_2 , and having ruled out the possibility that the CAT had a low difficulty level, the fact that the time spent taking the CAT is unrelated to student performance on it leads to two final considerations. The first concerns the very nature of the test: because it is taken online, students have the option of turning to others or the course manual itself for help answering the questions, enabling them to earn high grades. The second concerns the suitability of the time allotted to take the test, which could be considered excessive, as it allows students to access the aforementioned help.

4. CONCLUSIONS AND FUTURE RESEARCH

Students' participation in online CATs during a course has a dual positive effect on the results obtained subsequently with the final assessment test (final exam): it both increases the level of student participation in the exam and raises the grades students earn on it.

The assessment system used to determine the weight that online CATs will have in the final grade for a course has a significant influence on student performance, not only with regard to the CAT itself, but also the final exam and, thus, students' overall performance in the course.

Online CATs can be efficient tools both for self-assessment by students and as teaching tools to help students study and assimilate subject matter during the learning process. However, they are less useful as tools for the continuous assessment of students' knowledge, as their online nature makes it difficult to ensure the reliability of the results obtained on them.

Students' perceptions of these online CATs make them doubt their usefulness as assessment tools and thus have a considerable effect on the rate of student participation in them.

The design of a suitable assessment system for CATs is essential, not only to ensure that the educational and assessment goals set for the CAT itself are met, but also to encourage and increase participation and overall performance in the course.

The design of a CAT that is both suitable and rigorous in terms of characteristics such as the time allotted for its completion, systems for preventing data transfer, and the difficulty level of the questions is a key factor affecting the overall results obtained in the course, especially at universities whose methods are based on distance learning.

In light of the results and conclusions of this initial stage of research, a second stage is needed both to examine students' true motivations for voluntarily participating in these CATs more closely and to investigate the different types (in terms of their design) of online CATs being used in other courses and the results they are having.

AUTHOR INFORMATION

María Dolores Reina-Paz holds a PhD in business from the Universidad Nacional de Educación a Distancia (UNED, Spain). She is a professor of economics in the Business and Accounting Department of the Faculty of Economics and Business Administration at the UNED. Her research interests include marketing, education and tourism. Phone: +34 649 946 034; fax: +34 91 398 7355. E-mail: mreina@cee.uned.es

Ainhoa Rodríguez-Oromendía holds a PhD in business from the UNED. She is a professor of economics in the Business and Accounting Department of the Faculty of Economics and Business Administration, UNED. Her research interests include marketing, education and tourism. Phone +34 606 243 177; fax +34 91 398 6374. E-mail: arodriguez@cee.uned.es

Claudia Sevilla-Sevilla holds a degree in geography and history from the University of Barcelona (UB). She is an associate professor of economics in the Business and Accounting Department of the Faculty of Economics and Business Administration at the UNED. Her research interests include marketing, education and tourism. Phone: +34 675 897 190; fax: +34 91 398 7355. E-mail: csevilla@cee.uned.es (Corresponding author)

REFERENCES

1. Ávila Cañadas, M., Zaitegui De Miguel, N., Castillo Arredondo, I., Cantón Mayo, I., & Calatayud Salom, M. A. (2010). *La evaluación como proceso sistemático para la mejora educativa*. Ministerio de Educación.
2. Barberà Gregori, E., & de Martín Rojo, E. (2009). *Portfolio electrónico: Aprender a evaluar el aprendizaje*. Barcelona: Editorial UOC.
3. Cabra Torres, F. (2008). La calidad de la evaluación de los estudiantes: Un análisis desde estándares profesionales. *Magis: Revista Internacional de Investigación en Educación*, 1(1), 95-112.
4. Delgado García, A. M. (2005). *Competencias y diseño de la evaluación continua y final en el Espacio Europeo de Educación Superior* (Programa de Estudios y Análisis). Barcelona: Ministerio de Educación y Ciencia.
5. 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.
6. García González, G. (2004). *Retos en la práctica del proceso evaluativo en la educación superior no presencial*. Barcelona: Universitat Oberta de Catalunya.
7. López, V., Martínez, L., & Julián, J. (2007). La red de evaluación formativa, docencia universitaria y Espacio Europeo de Educación Superior (EEES). Presentación del proyecto, grado de desarrollo y primeros resultados. *RED-U. Revista de Docencia Universitaria*, 1(2), 2-19.
8. Margalef, L., Pérez, A., Urquizu, C., & Honduvilla, N. (2007). *Experiencias de innovación docente en la Universidad de Alcalá*. Alcalá de Henares: Universidad de Alcalá. Servicio de publicaciones.
9. Montoro, L. et al. (2011). Evaluación continua y autogestión del aprendizaje: Una visión interdisciplinar. *Univest 2011*, June 2011, 1.
10. Mosqueira-Rey, E. (2010). La evaluación continua y la autoevaluación en el marco de la enseñanza de la programación orientada a objetos (pp. 223-230). In: *Actas de las XVI Jornadas de Enseñanza Universitaria de la Informática (JENUI2010)*. Santiago de Compostela.
11. Ruiz Conde, E., & Parreño Selva, J. (2010). *La evaluación continúa de los aprendizajes: Una experiencia en grupos grandes*. Alicante: Facultad de Ciencias Económicas y Empresariales. Universidad de Alicante.
12. Vargas Quintero, M. V. (2008). Sentido pedagógico de la evaluación. Estudio de caso en la Facultad de Ciencias de la Pontificia Universidad Javeriana. *Universitas Scientiarum*, 13(3), 252-257.