



The relationship between formative assessment systems, academic performance and teacher and student workloads in higher education / Relación entre sistemas de evaluación formativa, rendimiento académico y carga de trabajo del profesor y del alumno en la docencia universitaria

Rosario Romero-Martín, Antonio Fraile-Aranda, Víctor-Manuel López-Pastor & Francisco-Javier Castejón-Oliva

To cite this article: Rosario Romero-Martín, Antonio Fraile-Aranda, Víctor-Manuel López-Pastor & Francisco-Javier Castejón-Oliva (2014) The relationship between formative assessment systems, academic performance and teacher and student workloads in higher education / Relación entre sistemas de evaluación formativa, rendimiento académico y carga de trabajo del profesor y del alumno en la docencia universitaria, *Infancia y Aprendizaje*, 37:2, 310-341, DOI: [10.1080/02103702.2014.918818](https://doi.org/10.1080/02103702.2014.918818)

To link to this article: <https://doi.org/10.1080/02103702.2014.918818>



Published online: 12 Jun 2014.



Submit your article to this journal [↗](#)



Article views: 603



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)

The relationship between formative assessment systems, academic performance and teacher and student workloads in higher education / *Relación entre sistemas de evaluación formativa, rendimiento académico y carga de trabajo del profesor y del alumno en la docencia universitaria*

Rosario Romero-Martín^a, Antonio Fraile-Aranda^b, Víctor-Manuel López-Pastor^b,
and Francisco-Javier Castejón-Oliva^c

^aUniversidad de Zaragoza; ^bUniversidad de Valladolid; ^cUniversidad Autónoma de Madrid

(Received 19 January 2013; accepted 22 March 2013)

Abstract: The purpose of this study was to analyse the extent to which the implementation of formative assessment systems in higher education influences students' academic performance; student and teacher workloads; and student satisfaction with such processes. Data come from 39 cases that apply formative assessment systems in higher education, which correspond to 15 Spanish universities and cover a total of 30 different tutors and 3625 students. Descriptive statistics and ANOVAS have been applied. The results suggest that the development of formative assessment systems: (a) produces a significant improvement in the academic performance of students; (b) is a workload that can be managed within the academic work set, both for students and teachers; and (c) that students show high levels of satisfaction with the formative assessment systems in which they participate and consider that they learn better with these as opposed to more traditional models.

Keywords: formative assessment; higher education; academic performance; student and teacher workloads; student satisfaction

Resumen: El objetivo del presente estudio fue analizar la influencia que tiene la aplicación de sistemas de evaluación formativa en la docencia universitaria en el rendimiento académico del alumnado, en la carga de trabajo de alumnado y profesorado y en la satisfacción del alumnado en dichos procesos. Los datos proceden de 39 casos que emplean sistemas de evaluación formativa en la docencia universitaria, que corresponden con 15 universidades españolas y que abarcan un total de 30 asignaturas diferentes y 3625 alumnos. Se aplican estadísticos descriptivos y ANOVAS. Los resultados parecen indicar que el desarrollo de sistemas de evaluación formativa: (a) produce una considerable

English version: pp. 310–324 / *Versión en español:* pp. 325–339

References / *Referencias:* pp. 339–341

Translation / *Traducción:* Julie Waddington

Author's Address / *Correspondencia con los autores:* Francisco-Javier Castejón-Oliva, Facultad de Formación de Profesorado y Educación, Universidad Autónoma de Madrid, C/Francisco Tomás y Valiente, 3, 28049 Madrid, España. E-mail: javier.castejon@uam.es

mejora en el rendimiento académico del alumnado; (b) es una carga de trabajo asumible dentro de la dedicación académica establecida, tanto para el alumnado como para el profesorado; (c) el alumnado muestra una alta satisfacción con el proceso de evaluación formativa en el que participa, percibiendo un mayor aprendizaje respecto a modelos más tradicionales.

Palabras clave: evaluación formativa; educación superior; rendimiento académico; carga de trabajo; satisfacción del alumnado

In education, assessment acts like a lens through which we can obtain information on how the formative process is evolving, enabling us to check if the process is or is not meeting our expectations and plans. Just over a decade ago, while already immersed in the process of convergence towards the European Higher Education Area (EHEA), Zabalza (2001) was already questioning the disconnection between assessment and educational activities in universities: 'The big problem with assessment is that it has ended up being separated from the formative process as if it were an isolated or self-sufficient piece of the curricular structure' (p. 271). In his review of the current state of play, he referred to a study carried out at the end of the 1990s (Fernández-Pérez, 1989) and to another from a decade later (Trillo & Porto, 1999), in which it is shown that higher education assessment processes present serious deficiencies and that, in order for these to be improved, the following points should be taken into account: (a) the greater use of continuous and formative assessment as opposed to piecemeal and stand-alone assessments; and (b) to consider assessment as a process for helping and guiding the student in their learning.

These arguments reflect a question concerning how to reorient assessment as a formative process in which students have a greater awareness of how they are developing as an alternative to summative approaches.

More recent studies (Gutiérrez, Pérez-Pueyo, Pérez-Gutiérrez, & Palacios, 2011; Palacios & López-Pastor, 2013) seem to indicate that the situation is improving in Spanish universities as a result of the gradual increase in the percentage of teachers using formative assessment systems (FA). The results show that, although still in the minority, this does represent an upward trend and that this change has developed more clearly than in previous decades, especially in the practical implementation of the new titles given to degree and Master's programmes (López-Pastor, Manrique, & Vallés, 2011).

The immersion of our universities into the 21st century, in the midst of the convergence towards EHEA, has led to a Copernican revolution in the formative paradigms of the university. The focus is no longer on the tutor, but on the student and the attainment of competences (Ministry of Education, 2006): a focus according to which tutors have to make these competences visible to society (Torrance, 2012).

With regard to assessment in higher education, López-Pastor (2009) supports the development and use of FA and shared systems whose aim is to improve student learning, the work of the tutor, and the teaching/learning process; a form of assessment that requires student participation in the process as a means of

improving their learning and the development of continuous and autonomous learning competences. Similar approaches can also be found in the studies carried out by Boud and Falchikov (2007); Brown and Glasner (2003), and by Falchikov (2005).

We are seeing a growing interest in the use of FA as a way of improving the quality of higher education within ENQU (The European Association for Quality Assurance in Higher Education), and of the quality of universities in general. In Spain, we have seen a proliferation of conferences, courses and scientific publications on FA in higher education, the contributions of which have addressed the following topics: (a) assessment of competences; (b) assessment instruments and techniques and their viability; (c) learning outcomes and the relation between FA and academic performance; (d) teacher and student workloads; (e) the use of ICT in FA; (f) teacher and student satisfaction with the system; (g) good FA practices in higher education; and (h) the relationship between teaching methodology and FA. For studies on these topics we can consult Arribas (2012), Buscá, Pintor, Martínez, and Peire (2010), Castejón, López-Pastor, Julián, and Zaragoza (2011), Fraile (2010, 2012), Greenwald and Gillmore (1997), Julián, Zaragoza, Castejón, and López-Pastor (2010), López-Pastor et al. (2011), Santos, Martínez, and López-Pastor (2009) and Zaragoza, Luis-Pascual, and Manrique (2008).

In this article, we are going to focus specifically on three of these: (a) student satisfaction with the development of FA Systems in higher education; (b) the relationship between the use of these systems and the students' academic performance; and (c) what the use of FA systems entails in terms of workload for students and teachers. We need to bear in mind, however, that FA means much more than the mere application of the appropriate instruments; it involves a redefining of the concept of assessment, with a focus more on the processes rather than on the functions of the FA itself (Taras, 2009).

Advantages of formative assessment in higher education

Like any other kind of educational intervention, the application of FA can have a series of advantages and disadvantages. One of the primary advantages of FA comes from the definition of the term itself: that the information obtained is used to reorient the process and, consequently, to take decisions about steps needed to improve it (Torrance, 2012). Although knowledge of the results of this process could be perceived as an unbridgeable gap by students — the distance between what they know, or what they know how to do, and what they ought to know (Taras, 2005; Torrance, 2012) — this difficulty can be overcome if they are reoriented in the right direction, since help must also be given on how to improve the work in order to attain the level required (Torrance & Pryor, 2001).

Equally, FA also provides teachers with insight into their own progress, helping them to find better alternatives. If one of the goals of higher education is to help students become 'independent and critical learners' (Torrance, 2012), the advantage of this kind of assessment is that it facilitates this general and transversal competence. In addition, students have a certain amount of control over

their learning which increases their competency and autonomy (Rué, 2009). Thus, students seem to indicate that there is a link between FA and responsibility, involvement in and control over the students' own learning (Boud & Falchikov, 2007; Falchikov, 2005).

Taras (2005) points out that the benefits of using FA for teachers include being able to discuss and clarify the assessment criteria and inform the students about what they need to reflect on. The work carried out by López-Pastor, Pintor, Muros, and Webb (2013) on the development of FA systems in higher education shows that: (a) FA has a positive impact on student participation and their learning; and (b) leads to a reduction in the numbers of students dropping out or not attending.

Disadvantages of formative assessment in higher education workload and other concerns

The literature shows one of the main disadvantages of the FA system to be the increase in time and work it involves for teachers and students (Bennet, 2011; Gibbs & Simpson, 2004–05). Nevertheless, in Julián et al. (2010) and López-Pastor et al. (2013), studies are carried out on the workload that the development of FA processes in higher education entails and the results show that these processes are entirely viable for both students and teachers.

Other disadvantages are also highlighted. Winchester and Winchester (2012) studied assessment in a virtual environment and found that, after an initial stage in which students were highly motivated, their motivation dropped when the tasks turned into a routine and started to require a significant amount of time, meaning that the students' response was ultimately superficial. Torrance (2012) thinks that there is a dominant belief within the current field of education in relation to the effectiveness of giving students formative feedback; however, the scientific research on the impact of FA seems to be insufficient to date. Therefore, in line with Dunn and Mulvenon's (2009) work, even though a good number of studies exist on the benefits of the application of this kind of assessment, some methodological problems make it impossible to extend the conclusions on a general basis. Other authors think that we are in difficult times in which we need to agree upon a lexicon that can be shared by all educational agents, including teachers, researchers, the politicians responsible and parents and students, so as to promote more productive communication to ensure the maximum benefits for all those involved. In this regard, Black, Harrison, Hogden, Marshall, and William (2005) are of the opinion that FA does not necessarily improve students' achievements in all situations, with all teachers, or on all occasions, but rather that, in general, while FA is an effective intervention tool, its difficulties should not be underestimated.

Formative evaluation and academic performance of students

With regard to the use of FA systems and processes and students' academic performance, in a case study on the use of FA and shared processes in initial

teacher training, López-Pastor (2008) finds a high level of academic performance both in terms of the percentage of students who pass the subject (94%) as well as in terms of the grades attained, which show a high level of academic achievement for most students. In a study on the development of FA systems in higher education with a sample of 1786 students from 10 universities, Zaragoza et al. find that the total percentage of students passing is 67.2% of the total, with 17.6% failing and 15.2% not submitting. Moreover, 50.5% attained grades equal to or higher than very good (2:1). Significant differences are also observed in terms of academic results; from cases in which 75–83% of students attain grades equal to or higher than very good, to those in which there are very high levels of non-submission (37–64%), that tend to be related to subjects in which the majority of students have opted to take a final exam instead of following the continuous FA route, or in cases in which the requirements of the FA route are very low and the weight of the final exam is high, meaning that most students continue to follow the same habits that are customary with more traditional systems (to study only for the final exam). Arribas (2012) analyses the effects of the assessment and qualification system used in the academic performance of 2192 students from 14 Spanish universities. The results confirm that the FA system produces the best results both in terms of success rates as well as in performance.

Hence, the development of FA systems tends to improve students' academic performance and their evaluation of the subject. This concurs with Buscá et al. (2010), in a study of 2476 students from 14 universities: when students are asked to choose the learning and assessment route, most of them choose the FA and continuous route (87%). Arribas (2012) and López-Pastor et al. (2013) find similar results with large sample sizes of students from different university programmes.

Another issue analysed is the different levels of academic performance achieved by students depending on the kind of learning and assessment route chosen (FA and continuous versus final exam or the mixed route). López-Pastor (2008) finds significant differences between both routes; students who opt for FA processes or mixed systems obtain a higher percentage of the high marks, while those who follow the final exam route show much lower levels of achievement and the majority of fails and non-submissions. This coincides with the arguments advanced by Black and Wiliam (1998) that academic performance in students following FA processes is better than in those who opt for other forms of assessment. Different authors agree on the strong influence that FA can have on learning and on students' academic performance (Brown & Glasner, 2003; López-Pastor, 2009). In Gibbs (2003), for example, we can find a case study in which the application of FA and shared systems in engineering meant that the percentage of passes increased from 45% to 75%. In contrast, in Castejón et al. (2011), we find significant differences in academic performance between the FA cases in which students have to carry out activities and produce documents in line with fixed deadlines, and cases in which no such deadlines exist and in which students can submit the documents freely at any time during the term. The results seem to indicate that it is important to fix specific deadlines for carrying out activities and

submitting documents in FA systems in order to get the most out of the learning process, which, in turn, tends to generate greater academic achievement.

On the other hand, according to Hernández (2012), students were aware that the FA could involve an extra workload thus generating anxiety about the marks they would obtain. Buscá et al. (2010) indicate similar results on finding that, in some cases, FA systems seem to generate reticence and concerns among students about their academic performance. By contrast, other cases of the same study show that the development of FA processes may also generate worrying levels of confidence in the certainty of success, which can even end up determining the final grade attained.

A final topic of interest on this subject is highlighted by Dunn and Mulvenon (2009), who point to the improvements produced among low-performing students as one of the most positive aspects of FA.

In view of the findings shown, our aim in this study is to analyse the influence that the application of FA systems has in higher education on student satisfaction with these systems, on improvements in student performance, and on teacher and student workloads.

Method

Design

This study is part of the first cycle of action-research within an R&D project whose aim is to develop formative and shared assessment systems in higher education through a multiple case analysis (Stake, 2006) that, while not intending to present itself as a comparative study, establishes common elements among the different cases, thus enabling an analysis to be made of a set of data taken from similar processes that may provide generalizable results. Each case focuses on a particular subject, usually on a termly basis, in relation to formative and shared assessment. In each case, information has been gathered systematically, through an instrument drafted on an 'ad hoc' basis, on the following aspects: student and teacher workloads, academic performance, satisfaction and learning. With all this information in hand, the teachers involved in the study completed detailed and structured reports on each experience, presenting alternative routes for the different processes of learning, of assessment and of grading the students' achievements for the subject. Three types of situation emerge from the different cases, depending on the routes offered by the teacher: (a) continuous route only; (b) continuous route and final route; and (c) continuous route, final route and mixed route.

The continuous route is based on a formative and continuous assessment process throughout the whole term, with no final exam being necessary or, if there is one, with the results only contributing to a part of the final mark; even though students must fulfil a series of requirements (continuous attendance, submission of work on time and in the right way, correction of the errors indicated by the tutor, achievement of the minimum standards set for the assessment activities...). The most substantial part of the final mark comes from the learning activities carried out throughout the term which are linked to processes in which

Table 1. Assessment routes offered by the teacher for the case study.

Assessment routes offered	Frequency	Percentage
Continuous route only	6	15.3
Continuous route and final exam	26	66.6
Continuous route, final exam and mixed route	7	17.9

students have participated in the assessment (self-assessments, peer assessments and shared assessments), although not in all cases.

The final route involves a final exam based on one or various tests (theoretical, practical, project-based, etc.). The mark is awarded on the basis of these tests. In order to pass, the minimum standards need to be achieved for each test. Students are not usually required to fulfil any other requirements during the term.

The mixed route is developed with a combination of the continuous and final methods. It is based on a final exam as well as learning activities carried out throughout the term, with more or less regular attendance being required, according to the grading criteria established for each subject. The final grade is the result of a combination of the final exam and the learning activities carried out. The requirements are less demanding than with the continuous route and could be, for example: attendance at over 50% of the sessions, development of a learning project with the support of a tutor, and passing a final theoretical exam.

In each case, tutors offer alternatives in the assessment and learning routes, and three types of case emerge: (a) continuous route only with 15.5%; (b) continuous route and final exam route, with 66.6%; (c) continuous route, final exam and mixed route with 17.9%. The distribution of cases according to the different options can be seen in [Table 1](#).

Population

The population of this study comprises 39 university teachers from 15 Spanish universities from most of Spain's autonomous communities. These teachers teach on a total of 30 different subjects of initial teacher training, with a total of 3625 students, representing a significant sample size. The breakdown of this data appears in [Table 2](#).

All the teachers and students have accessed the implementation of the study through a research project, with the corresponding permission to use the data having been granted.

Instruments

In each case, the teacher was asked to complete a structured report on the planning of and the results obtained from the application of a formative assessment system in a university subject, on the basis of the information gathered during the implementation of the system. The instrument was designed by a group of experts

Table 2. Population of study by participating teachers and students.

University	No. participating teachers	Percentage	No. participating students	Percentage
La Laguna	4	10.3	448	12.3
Valladolid	12	30.8	922	25.4
Autonomous University of Barcelona	2	5.1	250	7.0
Autonomous University of Madrid	2	5.1	116	3.2
Vic	6	15.4	378	10.4
León	2	5.1	193	5.3
Burgos	1	2.6	251	7.0
La Coruña	1	2.6	110	3.0
Zaragoza	2	5.1	115	3.1
Lleida	2	5.1	176	4.9
Alcalá	1	2.6	33	0.9
Murcia	1	2.6	48	1.3
Salamanca	1	2.6	90	2.5
Castilla la Mancha	1	2.6	124	3.4
Granada	1	2.6	371	10.2
Total	39	100.0	3625	100.0

with considerable experience in formative assessment in higher education and in this kind of study (Buscá et al., 2010; López-Pastor et al., 2013; Navarro, Santos, Buscá, Martínez, & Martínez, 2010; Zaragoza et al., 2008). The validation process was carried out during two different academic years until settling on the definitive instrument. After the initial design developed by the group of experts, it was then checked by a broad group of university teachers who were accustomed to using formative assessment systems in their subjects. Once the instrument had been improved, it was then put into practice during the course, after which it then went through another double process of analysis and improvement until an instrument which could be easily understood by the interviewees and which met the objectives of the study was attained. Subsequently, it was applied again in an experimental way during a second course and the results obtained were used to carry out a final cycle of analysis which helped to establish the definitive instrument to be used (see Table 3) (Manrique, Vallés, & Gea, 2012).

This consists of a set of four different sections: (a) a group of items in which the context is characterized by the different assessment approaches used by the teacher (type of subject, number of students, hours of class, tutorials, activities carried out by students, etc.) and in which the type of assessment method to be used is indicated, along with the different routes offered and the number of students who opt for each route; (b) a group of items in which satisfaction with the assessment processes followed is indicated. An anonymous questionnaire is used for students to evaluate the subject with a Likert scale of 5 points (0–4). During the validation process, the Cronbach alpha coefficient obtained was high

Table 3. Elements constituting the instrument where data related to the subject are included.

Aspect	Constitutive elements
Data related to context	University, Institution, Qualification, Course, Name of subject. Type of subject. Total teaching credits allocated to teacher. Type. Total class time per week for the teacher. Total no. of students
Assessment system details	Main competences that the subject develops and the relation with learning activities; with assessment activities; and with the criteria established for the grading of the subject
Learning and assessment routes	Single route system, characteristics, requirements and associated activities. System based on different kinds of routes: characteristics, requirements and associated activities
Teaching/learning activities and their relation with assessment processes	Types of activities and frequency of application
Application of assessment instruments and learning support	Assessment instruments. Student feedback using the assessment instrument. Explanation of how the formative process occurs with the instruments used.
Grading criteria, evidence of learning and assessment protocol	Main criteria for grading students work. Evidence of learning students asked to present. Assessment protocol of the evidence requested. How reflected in the grades.
Types of cognitive abilities required by the instruments and assessment techniques	Cognitive abilities. Instruments and techniques for assessing. Frequency of use.
Ethical criteria and nature of the assessment decisions made	Ethical principles
Development of assessment systems using ICT	Assessment of formative applicability
Advantages noted in the assessment system	Qualitative comments
Disadvantages perceived in the assessment system	Qualitative comments
Academic performance and assessment of it	Distribution of the number and percentage of students according to grades. Proportion and performance for each different route.
Teacher and student workload	Teacher: no. of hours per subject per month. Student: no. of hours per activities per month

(0.84), with a KMO index of 0.807, and a value of $p > .00$ for Bartlett's test of sphericity. It has now been used for three academic years and is carried out anonymously with all students in the last class of the subject; (c) the scores

obtained by each student, organized by partial or total scores, according to the learning route followed. The data used is taken from the official grades recorded for the subject; and (d) a group of items which indicates the workload that corresponds to the teacher and the students.

Procedure and analysis

Each teacher undertakes to develop a formative assessment system in at least one of the subjects they teach during the academic year, whether this be in the first or second term. At the beginning of term, each teacher provides the information related to the design of the subject, corresponding to section (a) of the instrument. Throughout the term the students and teacher responsible for delivering the subject keep a record of the number of hours they dedicate to the teaching/learning and assessment activities, using ad hoc annotation instruments and workload calculations which are completed on a weekly basis. At the end of term, the instrument is completed with the academic results obtained and with the data collected throughout its application.

Once the teacher's structured report has been completed, the data is stored and analysed with the SPSS 18.0 programme. Descriptive statistics and ANOVAS are applied. The results for all the teachers belonging to the R&D project are pooled together in order to improve the assessment systems used, particularly in relation to the viability of the FA systems and ways of solving problems encountered when applying them.

Results and discussion

Learning and student satisfaction

Very high scores have been recorded by students in the items that refer to the learning achieved during FA processes, as well as to their satisfaction with these. [Table 4](#) shows the mean scores given by students to some aspects of the formative assessment process carried out by the teacher, using a Likert scale in which 4 is the maximum score.

Table 4. Mean of the scores given by students to the assessment system used by the teacher (Likert scale of 0 to 4).

Terms used to score learning	Mean
5.- Student practices active learning	3.42
6.- Student is more motivated, the learning process is more motivating	3.08
7.- Grading is fairer	3.13
8.- Improves academic tutelage (supervision and student support)	3.03
9.- Allows for functional learning	3.08
10.- Generates meaningful learning	3.41
11.- Students learn much more	3.32

These items are grouped into two sets: those that refer to learning (5, 9, 10 and 11) and satisfaction with the FA processes followed (items 6, 7 and 8). The students' evaluations of perceived learning can be considered very high, and especially high in some aspects such as 'achievement of active learning' (3.42), the generation of meaningful learning (3.41), and the high level of agreement with the suggestion that 'you learn much more' than with more traditional models (3.32). With regard to evaluations of their satisfaction with these kinds of learning and assessment processes, the scores obtained are also very high; particularly noteworthy is the item indicating a greater perception of fairness in the subject grading system. These results are consistent with those found by Dunn and Mulvenon (2009), López-Pastor et al. (2013) and Taras (2005), who document the improvement of educational results through the use of FA. On the other hand, after developing FA systems in universities, Álvarez, Grau, and Tortosa (2010) conclude that attendance at classes increases, as does student involvement in tasks, the quality of the tasks and projects carried out, and the academic results obtained. Hernández (2012) also finds that the kind of assessment that provides students with feedback to facilitate their learning is more effective and efficient.

Academic performance

As shown in Table 5, general performance can be seen as good, with the percentage of students passing the subject being fairly high (around 85%). The percentage of non-submissions is low (7.8%), as is the percentage of fails (7.6%). Most students achieve very good marks (2:1), followed by good (2:2) and distinctions (1st). Spanish universities tend to establish a norm whereby distinctions with honours can only be granted to a maximum of 5% of the students on any given subject.

With these data in hand, we have calculated an 'academic performance index', applying the formula used by universities for student academic records, which involves calculating the sum derived from multiplying the marks obtained on a scale of 0–4 (non-submissions and fails by 0, good by 1, very good by 2, distinction by 3 and distinction with honours by 4), and then dividing this sum by the number of subjects taken which provides an index of between 0 and 4. In this way, in addition to obtaining an index for overall performance as well as one by route followed, we have also been able to carry out ANOVAS to observe if there are any significant differences in performance depending on the different

Table 5. General performance in all cases, differentiated according to grades.

	Grades					
	NS	Fail	Good	Very Good	Distinction	Distinction with Honours
Percentages	7.8	7.6	23	47.9	11.8	3.66

Table 6. Index of academic performance according to the learning and assessment route chosen.

Route	Continuous	Mixed	Exam	Total
Performance index	1.91	1.48	.53	1.57
INTER-ROUTE ANOVA SCORES				
Sum of squares 29.98	<i>gl</i> 2	Mean square 7.419	<i>F</i> 29.346	Sig. .000

Table 7. Percentages of grades obtained by students for each route.

Route	Grade					
	NS	Fail	Good	Very Good	Distinction	Distinction with Honours
Continuous	2.6	2.1	10.8	28.7	7.0	2.54
Mixed	2.1	4.0	9.9	13.4	6.0	1.0
Final exam	5.0	4.9	10.11	1.1	–	–

routes offered. Table 6 shows the performance indexes according to the assessment route selected by the student, as well as the result of the inter-route ANOVA.

This table enables us to carry out two types of analysis: (a) there is a strong difference in academic performance between the different routes, which is statistically significant; and (b) the route showing the best performance is the continuous route, with a mean close to very good, followed by the mixed route, with a score slightly above good. The route showing the lowest academic performance is the final exam route, with scores clearly in the region of fail.

In order to analyse this difference between the different learning and assessment routes in more detail, Table 7 shows the distribution of student grades according to the route followed.

It can be observed that the final exam route includes the highest percentage of fails and non-submissions, as well as very low scores among those who pass the subject following this particular route. On the other hand, the majority of students who opt to follow the continuous route achieve high grades (very good, distinction or distinction with honours) and the percentage of insufficient results is very high, both in terms of non-submissions and fails. The mixed route has quite a higher percentage of passes than insufficients (29% as opposed to 6%), but achieves somewhat lower percentages in the high grades; interestingly, the most significant differences seen in contrast to the continuous route is for 'very good', while the percentages are very similar for distinctions and are close for distinctions with honours. These results are consistent with those found by Arribas (2012), López-Pastor (2008) and Zaragoza et al. (2008) on the good academic

Table 8. Average workload in hours for teachers distributed for the total of the study in each term.

Type of activity	Total 1 st term	Total 2 nd term
CP	50.9	48.7
CW	47.1	84.5
VC	24	24.6
Total	122	158

(CP = class preparation; CW = correcting work; VC = virtual campus).

performance usually obtained when FA Systems are used in higher education. Similarly, the studies carried out by Buscá et al. (2010), Castejón et al. (2011), López-Pastor (2008, 2009), López-Pastor et al. (2011, 2013) and Santos et al. (2009) also find differences in performance levels between the different learning and assessment routes selected by students, in the sense that students who follow the FA route attain higher levels of academic performance than students who follow the final exam route.

Teacher workloads

Table 8 shows the results of teacher workloads for those using formative assessment methods.

The mean for hours spent preparing classes (CP) is higher, as is the time dedicated to ‘correcting student work’ (CW). The lowest means observed are for the time dedicated to ‘using and maintaining the virtual campus’ (VC) (Moodle, WebCt, etc.) for developing material. The total time dedicated and the internal distribution of CW shows some variations between terms.

The mean workload is 122 hours for the first term and 158 for the second. If we divide this by the 15 weeks for each term, the mean result is between eight and 10 hours per week, which does not appear to point to an excessive workload. It should be noted that we are referring to means and that variations between cases may be considerable. The distribution of workload is accumulated particularly in the final months of each term. Hernández (2012) has highlighted teachers’ concerns over the high degree of work that the implementation of continuous FA systems entails, as a result of the need to give students effective feedback at the right time and bearing in mind the high numbers of students involved, in comparison with other more traditional systems.

Student workloads

The mean total student workload is approximately 85 hours per year, which means around 42.5 hours per term. The majority of subjects taken on a termly basis are allocated six credits and, with the current ECTS system, each credit is equivalent to an average number of 25 hours work for the student, of which 10 tend to be

Table 9. Workload for students (with an average of 87 hours per year/subject) according to the activity set by the teacher.

Time dedicated to:	Total
Lectures and reviews	24.1%
Preparation and application of practical sessions	23.0%
Tasks and reports on specific subjects	33.3%
Reviewing and correcting work (in groups and in tutorials)	19.5%

class-based and 15 used for independent study by the student. Therefore, a subject of six credits means a total of 90 hours of independent study by the student. Since the mean workload shown is 42.5 hours, it could be argued that the development of FA Systems in higher education does not involve an excessive workload for students, but is, rather, clearly within the expectations set by the ECTS system.

We have also analysed the distribution of student workload, organized according to the learning activities carried out during the subject (see Table 9). The data are presented in mean percentages in relation to the total time spent.

The highest percentage is shown for carrying out projects and reports on class topics (a third of the total workload). This is followed by lectures and reviews and the preparation and application of sessions; leaving the workload related to reviewing and correcting submitted work in final position.

This workload distributed by months (Table 10) leads us to confirm that the greatest weight of the students' workload always comes in the last month of the term, while the lightest weighting is found in the first month; in the intermediate months it is distributed fairly evenly.

Consistent with these results, López-Pastor et al. (2013) observe certain disparities between the subjective opinions of students and teachers and the actual data on workloads which are collected systematically and within the officially established benchmarks. For these authors, FA systems involve much higher workloads than traditional final exam systems, where students limited themselves to studying during the last month and the teacher simply corrected the exams. The perception of teachers and students can thus be explained more by the comparison they make with traditional models than by the evidence found on the actual workloads involved. Similar results are found by Julián et al. (2010) who indicate that the application of FA in higher education is perfectly viable within the workloads established in the ECTS system for students and teachers.

Table 10. Student workload distributed by month and term (in hours).

	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Time spent on work	10.1	19.5	19.2	22.7	34.1	10.8	13.5	12.8	22.3

Conclusions

Our aim in this study has been to analyse the influence that the application of FA systems has in higher education on student satisfaction with these systems, on improvements in student performance, and on teacher and students workloads.

The results show very high scores have been recorded by students in the items that refer to the learning achieved during FA processes, as well as to their satisfaction with these.

With regard to academic performance, statistically significant differences can be observed between the different learning and assessment routes, with the FA and continuous route obtaining higher performance levels, followed by the mixed route. The final exam route shows very low performance levels.

In terms of teacher workloads, the mean result is between eight and 10 hours per week of non-contact time (class preparation, correcting student work and student activities, maintaining web and virtual tutorials with students). The third activity (web/virtual-based activities) is the one that seems to require the least time.

Finally, the mean workload shown for students is 42.5 hours per subject taken per term (six credits), clearly within the expectations set by the ECTS system, which means that it could be argued that the development of FA Systems in higher education does not involve an excessive workload for students.

With regard to the limitations of the study, we have observed that allowing students to choose from different assessment routes means that the majority of good students tend to choose the continuous route, therefore, this variable may be affecting the final result. This is a difficult problem to solve in view of the difficulties involved in implementing other design types (experimental and quasi-experimental). A less significant limitation relates to the sample size, in the sense that it would be interesting to accumulate the data with larger populations.

The results show the advantages of advancing further in the development of FA systems in higher education, bearing in mind the considerable improvements that these systems represent for academic performance, learning and student satisfaction. On the other hand, they do not seem to involve an excessive workload for students or teachers as is often thought; consequently, we suggest that formative assessment processes and instruments should be widely applied in higher education, in such a way that such systems do not represent an optional assessment system in teaching guides, but are rather an approach promoted and backed up by academic institutions.

Relación entre sistemas de evaluación formativa, rendimiento académico y carga de trabajo del profesor y del alumno en la docencia universitaria

La evaluación es para la educación como una lente a través de la cual obtener información de cómo evoluciona el proceso formativo, permitiéndonos comprobar si dicho proceso se ajusta o no a lo previsto y programado. Hace poco más de una década, ya inmersos en el proceso de Convergencia hacia el Espacio Europeo de Educación Superior (EEES), Zabalza (2001) cuestionaba la desconexión entre la evaluación y la actividad formativa en la universidad: ‘El gran drama de la evaluación habitual es que ha acabado independizándose del proceso formativo como una pieza aislada y autosuficiente de la estructura curricular’ (p. 271). En su revisión del estado de la cuestión hacía referencia a un estudio de finales de los noventa (Fernández-Pérez, 1989) y otro de una década después (Trillo y Porto, 1999), en los que manifiesta que la evaluación en la enseñanza universitaria presentaba graves deficiencias y que, de cara a mejorar su calidad, deberían tenerse en cuenta: (a) un mayor uso de la evaluación continua y formativa frente a evaluación puntual y aislada; y (b) plantear la evaluación como un proceso de ayuda y orientación al aprendizaje del alumnado.

Esos argumentos reflejan una demanda sobre cómo reorientar la evaluación como proceso formativo, en el que el alumno tuviera un mejor conocimiento de sus avances y como alternativa a una evaluación de carácter finalista.

Trabajos más recientes (Gutiérrez, Pérez-Pueyo, Pérez- Gutiérrez, y Palacios, 2011; Palacios y López-Pastor, 2013) parecen indicar que esta situación está mejorando en la universidad española, al incrementarse poco a poco el porcentaje de profesorado que utiliza sistemas de evaluación formativa (EF). Los resultados muestran que, aunque son una minoría, se trata de una tendencia al alza y que el cambio ha evolucionado de una forma más clara que en décadas anteriores, especialmente en la implementación práctica de los nuevos títulos de Grado y Máster (López-Pastor, Manrique, y Vallés, 2011).

La inmersión de nuestras universidades en el siglo XXI, en pleno proceso de convergencia hacia el EEES, ha conllevado un giro copernicano en los paradigmas formativos universitarios. El foco deja de estar en el profesor para centrarse en el alumnado y en la consecución de sus competencias (MEC, 2006), donde los formadores tienen que hacerlas visibles a la sociedad (Torrance, 2012).

Respecto a la evaluación en educación superior, López-Pastor (2009) defiende el desarrollo y utilización de sistemas de EF y compartida, cuya finalidad es mejorar el aprendizaje del alumnado, la labor del profesorado y del proceso de enseñanza-aprendizaje. Una evaluación que busca la participación del alumnado

en dichos procesos, como medio de mejora de su aprendizaje y el desarrollo de competencias de aprendizaje continuo y autónomo. Similares planteamientos pueden encontrarse también en las recopilaciones realizadas por Boud y Falchikov (2007), Brown y Glasner (2003), y por Falchikov (2005).

Estamos ante un creciente interés por el uso de la EF como mejora de la calidad docente universitaria dentro de la ENQA (*The European Association for Quality Assurance in Higher Education*), y de la calidad universitaria en general. En España asistimos a la proliferación de congresos, cursos y publicaciones científicas en relación a la EF en educación superior, cuyas aportaciones atienden a los siguientes tópicos: (a) evaluación de competencias; (b) instrumentos y técnicas de evaluación y su viabilidad; (c) resultados de aprendizaje y relación entre EF y rendimiento académico; (d) carga de trabajo para el profesorado y el alumnado; (e) uso de las TIC en la EF; (f) satisfacción de profesores y alumnos con el sistema; (g) buenas prácticas de EF en educación superior; y (h) relación entre metodología didáctica y EF. Sobre estos tópicos podemos dirigirnos a Arribas (2012), Buscá, Pintor, Martínez, y Peire (2010), Castejón, López-Pastor, Julián, y Zaragoza (2011), Fraile (2010, 2012), Greenwald y Gillmore (1997), Julián, Zaragoza, Castejón, y López-Pastor (2010), López-Pastor *et al.* (2011), Santos, Martínez, y López-Pastor (2009) y Zaragoza, Luis-Pascual, y Manrique (2008).

En este artículo, vamos a centrarnos de forma específica en tres de ellas: (a) satisfacción del alumnado con el desarrollo de sistemas de EF en docencia universitaria; (b) relación entre el uso de dichos sistemas y el rendimiento académico del alumnado; y (c) qué carga de trabajo supone para alumnado y profesorado la utilización de sistemas de EF. Pero hemos de tener en cuenta que la EF implica mucho más que la mera aplicación de instrumentos adecuados, supone redefinir el concepto de evaluación, basándonos en los procesos y no tanto en las funciones de ésta (Taras, 2009).

Ventajas de la evaluación formativa en docencia universitaria

Como cualquier otro tipo de intervención didáctica, la aplicación de una EF puede tener una serie de ventajas e inconvenientes. Una primera ventaja de la EF parte de su propia definición: que la información que se obtenga sea utilizada para reorientar el proceso y, en consecuencia, para tomar decisiones sobre los próximos pasos a dar para mejorarlo (Torrance, 2012). Aunque el conocimiento de los resultados sobre ese proceso corra el riesgo de presentarse ante el alumno como una ‘brecha’ insalvable, la distancia entre lo que sabe o sabe hacer y lo tendría que saber (Taras, 2005; Torrance, 2012), una acertada reconducción puede salvar esa dificultad, ya que también se debe indicar cómo mejorar el trabajo para alcanzar el nivel requerido (Torrance y Pryor, 2001).

De igual manera, la EF también informa al profesor de su propio progreso ayudándole a encontrar mejores opciones. Si uno de los objetivos de la educación superior es que los estudiantes sean ‘aprendices independientes y críticos’ (Torrance, 2012), este tipo de evaluación presenta la ventaja de facilitar esa competencia general y transversal. También, con la EF el estudiante

debe tener un cierto control sobre su aprendizaje, lo cual incrementará su competencia y autonomía (Rué, 2009). Por tanto, los estudios parecen indicar que se establece un vínculo entre EF y responsabilidad, implicación y control sobre el propio aprendizaje de los estudiantes (Boud y Falchikov, 2007; Falchikov, 2005).

Taras (2005) indica los beneficios para el docente del uso de EF a partir de poder discutir y comprender los criterios e informar a los alumnos sobre qué reflexionar. El trabajo de López-Pastor, Pintor, Muros, y Webb (2013) sobre el desarrollo de sistemas de EF en educación superior, muestra que: (a) la EF tiene un impacto positivo sobre la participación del alumnado y su aprendizaje; y (b) que se produce una reducción de abandonos y faltas de asistencia entre el alumnado.

Inconvenientes de la evaluación formativa en docencia universitaria. La carga de trabajo y otras dudas

La literatura nos muestra como principal desventaja del sistema de EF el incremento de tiempo y dedicación que supone para el profesorado y el alumnado (Bennet, 2011; Gibbs y Simpson, 2004–05). No obstante, en Julián *et al.* (2010) y López-Pastor *et al.* (2013) se recogen estudios sobre la carga de trabajo que supone el desarrollo de procesos de EF en docencia universitaria, y los resultados muestran que son perfectamente viables para alumnado y profesorado.

También aparecen otros inconvenientes. Winchester y Winchester (2012) estudiaron la evaluación en un entorno virtual, encontrando que tras una primera fase de alta motivación del alumno, ésta bajó al hacerse rutinarias las tareas y requerir mucho tiempo, lo que llevó finalmente a que los alumnos respondieran de manera superficial. Torrance (2012) considera que en el ámbito educativo actual existe una creencia dominante, referida a la eficacia de proporcionar a los estudiantes retroalimentación formativa; no obstante, la investigación científica sobre el impacto de la EF parece ser aún insuficiente y, en la línea marcada por Dunn y Mulvenon (2009), aun existiendo un buen número de trabajos sobre los beneficios de la aplicación de este tipo de evaluación, algunos problemas metodológicos impiden generalizar las conclusiones. Estos autores consideran que estamos en tiempos vulnerables, donde necesitamos consensuar un léxico compartido entre todos los agentes educativos, entre los profesores, investigadores, los responsables políticos, padres de familia y estudiantes, que haga que exista una comunicación más productiva para conseguir el máximo beneficio para todos los involucrados. En este sentido, Black, Harrison, Hogden, Marshall, y Wiliam (2005) consideran que la EF no supone la mejora del logro del estudiante en todas las situaciones, con todos los profesores o en todas las ocasiones, si no, más bien, que la EF en general es una intervención eficaz, si bien para su estudio no deberán subestimarse las dificultades.

Evaluación formativa y rendimiento académico del alumnado

Respecto al uso de sistemas y procesos de EF y el rendimiento académico del alumnado, López-Pastor (2008), en un estudio de caso sobre el uso de una EF y compartida en formación inicial del profesorado encuentra un rendimiento académico muy elevado, tanto en lo que respecta al porcentaje de alumnado que supera la materia (94%) como a la distribución de las calificaciones, que muestran un elevado rendimiento académico en la mayor parte del alumnado. Zaragoza *et al.* (2008) recoge en un estudio en docencia universitaria sobre el desarrollo de sistemas de EF, con una muestra de 1786 alumnos de 10 universidades, que el porcentaje global de alumnos aptos supone un 67.2% del total, con un 17.6% de suspensos y un 15.2% de no presentados. Además, el 50.5% ha obtenido una calificación igual o superior a notable. También se observa bastante diversidad de resultados académicos; desde casos donde el 75–83% del alumnado obtiene una calificación igual o superior a notable, hasta aquellos donde aparecen porcentajes muy elevados de no presentados (37–64%), que suelen corresponder con asignaturas donde la mayor parte del alumnado ha optado por seguir la vía del examen final, en vez de seguir la vía de la EF y continua, o bien en casos donde los requisitos de la vía de EF son muy bajos y el peso del examen final considerable, por lo que la mayoría del alumnado mantiene hábitos propios de sistemas más tradicionales (limitarse a estudiar de cara al examen final). Arribas (2012) analiza los efectos del sistema de evaluación y calificación empleado en el rendimiento académico de 2192 alumnos de 14 universidades españolas. Los resultados confirman que la EF es el sistema que propicia los mejores resultados, tanto en tasa de éxito como de rendimiento.

Por tanto, el desarrollo de sistemas de EF suele mejorar el rendimiento académico del alumnado y la valoración de la asignatura. Esto coincide con Buscá *et al.* (2010), en un estudio con 2476 estudiantes de 14 universidades: cuando se da a elegir al alumnado la vía de aprendizaje y evaluación, la mayor parte elige la vía de EF y continua (87%). También lo recogen Arribas (2012) y López-Pastor *et al.* (2013), con muestras grandes de estudiantes de diferentes titulaciones universitarias.

Otro de los tópicos analizados es la diferencia de rendimiento académico en el alumnado en función de la vía de aprendizaje y evaluación elegida (EF y continua vs. examen final o vía mixta). López-Pastor (2008) encuentra fuertes diferencias entre ambas vías; el alumnado que opta por los procesos de EF o mixta acumula un mayor porcentaje de puntuaciones altas, mientras que los que siguen la vía del examen final muestran un rendimiento mucho más bajo y la mayoría de los suspensos y no presentados. Esto coincide con lo señalado por Black y Wiliam (1998), de que el rendimiento académico en los alumnos sometidos a procesos de EF es mejor que en aquellos que optan por otros tipos de evaluación. Diferentes autores coinciden sobre la fuerte influencia que la EF puede tener en el aprendizaje y en el rendimiento académico del alumnado (Brown y Glasner, 2003; López-Pastor, 2009). Por ejemplo en Gibbs (2003) puede encontrarse un estudio de caso en que la aplicación de sistemas de EF y compartida en ingeniería hizo que el porcentaje de aptos pasara de un 45% a un 75%. En cambio, en Castejón

et al. (2011) se encuentran fuertes diferencias en el rendimiento académico entre los casos de EF en los que el alumno tiene que realizar las actividades y documentos ajustados a plazos temporales fijos y los casos en que no existen plazos y el alumno puede entregar libremente dichos documentos a lo largo del cuatrimestre. Los resultados parecen indicar que en los sistemas de EF es importante fijar plazos concretos de realización de actividades y entrega de documentos para obtener un mayor provecho en el aprendizaje, que suele generar un mejor rendimiento académico.

Por otra parte, para Hernández (2012) los estudiantes eran conscientes de que la EF les podría suponer una sobrecarga y, con ello, propiciar ansiedad acerca de las calificaciones obtenidas. Similares resultados encuentran Buscá *et al.* (2010), al comprobar que, en algunos casos, los sistemas de EF parecen generar reticencias y dudas al alumnado sobre su rendimiento académico. En cambio, en otros casos del mismo estudio encuentran que el desarrollo de procesos de EF también puede generar una peligrosa confianza en la seguridad de éxito, lo que condiciona hasta cierto punto la calificación final.

Un último tópico de interés sobre esta temática es el señalado por Dunn y Mulvenon (2009), que apuntan como aspectos positivos de la EF las mejoras producidas sobre todo en los logros de los estudiantes de bajo rendimiento.

Debido a estas apreciaciones que se han mostrado, en el presente estudio tratamos de analizar la influencia que tiene la aplicación de sistemas de EF en la docencia universitaria en la satisfacción del alumnado con los mismos, en la mejora del rendimiento académico del alumnado y en la carga de trabajo de alumnado y profesorado.

Método

Diseño

Este estudio proviene de la realización del primer ciclo de investigación-acción, dentro de un proyecto de *I + D*, cuya finalidad es desarrollar sistemas de evaluación formativa y compartida en docencia universitaria mediante un estudio multicases (Stake, 2006), que sin pretender llegar a un estudio comparativo, se establecen elementos comunes entre todos los casos que permiten explorar un conjunto de datos provenientes de procesos similares, que puede llegar a resultados generalizables. Cada caso se centra en una asignatura, habitualmente cuatrimestral, desde procesos de evaluación formativa y compartida. En cada caso se ha recogido información de forma sistemática, mediante un instrumento elaborado 'ad hoc', sobre los siguientes aspectos: carga de trabajo para alumnado y profesorado, rendimiento académico, satisfacción y aprendizaje. Con toda esa información, el profesorado implicado en el estudio ha realizado informes detallados y estructurados sobre cada experiencia, presentando vías alternativas para los procesos de aprendizaje, evaluación y calificación de la asignatura. En los diferentes casos surgen tres tipos de situaciones, en función de las vías que oferta el profesorado: (a) sólo vía continua; (b) vía continua y vía final; y (c) vía continua, vía final y vía mixta.

La vía continua parte de un proceso de evaluación formativa y continua a lo largo de todo el cuatrimestre, no siendo necesario un examen final o, si se realiza, sólo corresponde con una parte de la calificación. Aunque, el alumnado debe cumplir una serie de requisitos (asistencia continuada, entrega de documentos en plazo y forma, corrección de los errores señalados por el profesor, superar los estándares mínimos establecidos para las actividades de evaluación,...). La mayor parte de la calificación final proviene del resto de actividades de aprendizaje realizadas durante el cuatrimestre, asociada a procesos de participación del alumnado en la evaluación (autoevaluación, coevaluación y evaluación compartida), aunque no en todos los casos.

La vía final consiste en un examen final, a partir de una o en varias pruebas (teórica, práctica, proyecto, etc.). La calificación se obtiene de dichas pruebas finales. Para conseguir una calificación positiva es preciso superar los estándares mínimos en cada prueba. No suele ser necesario cumplir ningún requisito durante el cuatrimestre.

La vía mixta se desarrolla entre medias de la continua y la final. Está basada en la realización de una prueba final y de actividades de aprendizaje durante el cuatrimestre, siendo necesaria una asistencia más o menos regular, en función de lo establecido en los criterios de calificación de cada asignatura. La calificación final proviene de una combinación entre las pruebas finales y las actividades de aprendizaje realizadas. Los requisitos son menos exigentes que la vía continua; por ejemplo, podrían ser: asistencia a más del 50% de las sesiones, realización de un proyecto de aprendizaje tutelado y la superación de un examen teórico final.

En cada caso, el profesorado oferta alternativas en las vías de evaluación y aprendizaje y aparecen tres tipos de casos: (a) sólo vía continua con un 15.5%; (b) vía continua y vía examen final, con un 66.6%; (c) vía continua, vía examen final y vía mixta con un 17.9%. La distribución de casos según dichas opciones puede comprobarse en la [Tabla 1](#).

Población

La población que ha participado en este estudio está formada por 39 profesores y profesoras, pertenecientes a 15 universidades españolas de la mayoría de las autonomías. Este profesorado abarca un total de 30 asignaturas de formación inicial. Con un total de 3625 alumnos, que representa una muestra considerable. El desglose de estos datos aparece en la [Tabla 2](#).

Tabla 1. Vías ofertadas en la evaluación que propone el profesorado para los casos de estudio.

Oferta de vías de evaluación	Frecuencia	Porcentaje
Sólo vía continua	6	15.3
Vía continua y vía examen final	26	66.6
Vía continua, vía examen final y vía mixta	7	17.9

Tabla 2. Población del estudio entre profesorado y alumnado participante.

Universidad	Nº profesorado participantes	Porcentaje	Nº alumnado participante	Porcentaje
La Laguna	4	10.3	448	12.3
Valladolid	12	30.8	922	25.4
Autónoma de Barcelona	2	5.1	250	7.0
Autónoma de Madrid	2	5.1	116	3.2
Vic	6	15.4	378	10.4
León	2	5.1	193	5.3
Burgos	1	2.6	251	7.0
La Coruña	1	2.6	110	3.0
Zaragoza	2	5.1	115	3.1
Lleida	2	5.1	176	4.9
Alcalá	1	2.6	33	0.9
Murcia	1	2.6	48	1.3
Salamanca	1	2.6	90	2.5
Castilla la Mancha	1	2.6	124	3.4
Granada	1	2.6	371	10.2
Total	39	100.0	3625	100.0

El conjunto de profesores y alumnos han accedido a la puesta en práctica del estudio a través de un proyecto de investigación, contando con los correspondientes permisos para exponer los datos.

Instrumentos

En cada caso, el docente elabora un informe estructurado sobre planificación y resultados de la aplicación de un sistema de evaluación formativa en una asignatura universitaria, a partir de los datos recogidos durante su puesta en práctica. El instrumento ha sido diseñado por un grupo especializado, con larga experiencia en evaluación formativa en docencia universitaria y en este tipo de estudios (Buscá *et al.*, 2010; López-Pastor *et al.*, 2013; Navarro, Santos, Buscá, Martínez, y Martínez, 2010; Zaragoza *et al.*, 2008). El proceso de validación se ha llevado durante dos cursos hasta elaborar el instrumento definitivo. Tras un diseño inicial desarrollado por el grupo de expertos, fue contrastado con un grupo amplio de profesorado universitario que utilizaba habitualmente sistemas de evaluación formativa en sus asignaturas. Una vez mejorado, se puso en práctica durante un curso, al final del cual volvió a someterse a un proceso doble de análisis y mejora hasta conseguir un instrumento con un buen nivel de comprensión para los encuestados y de adecuación a los objetivos del estudio. Posteriormente, se volvió a aplicar de forma experimental un segundo curso y con los resultados obtenidos se realizó el último ciclo de análisis, estableciendo con ello el instrumento definitivo (ver [Tabla 3](#)) (Manrique, Vallés, y Gea, 2012).

Tabla 3. Elementos que constituyen el instrumento donde se incluyen los datos relativos a la asignatura.

Dimensión	Elementos que la constituyes
Datos del contexto	Universidad. Centro. Titulación. Curso. Denominación de la asignatura. Tipo de Asignatura. Créditos totales de docencia del profesor. Tipología. Horas totales de clase a la semana para el docente. N° total de alumnos
Relaciones del sistema de evaluación	Principales competencias que desarrolla la asignatura y relación: con las actividades de aprendizaje; con las actividades de evaluación; y con los criterios de referencia para la calificación
Vías de aprendizaje y evaluación	Sistema vía única, características, requisitos y actividades asociadas. Sistema basado en varias vías: características, requisitos y actividades asociadas
Actividades de enseñanza-aprendizaje y su relación con la evaluación	Tipo de actividades y frecuencia de aplicación
Aplicación de instrumentos de evaluación y orientación del aprendizaje	Instrumentos de evaluación. Valoración hacia el alumno con la aplicación del instrumento. Explicación de cómo se logra la función formativa a través de los instrumentos empleados
Criterios de calificación, evidencias del aprendizaje y protocolos de valoración	Criterios de referencia para la calificación de la asignatura. Evidencias de aprendizaje solicitadas al alumnado. Protocolos de valoración de las evidencias solicitadas. Reflejo en la calificación
Tipo de capacidades cognitivas que requieren los instrumentos y técnicas de evaluación	Capacidades cognitivas. Instrumentos y técnicas para evaluar. Frecuencia de empleo
Criterios éticos y carácter de las decisiones de la evaluación	Principios éticos
Desarrollo del sistema de evaluación a través de las TIC	Valoración de la aplicabilidad formativa
Ventajas que se reconocen en el propio sistema de evaluación	Comentarios cualitativos
Inconvenientes que se perciben en el propio sistema de evaluación	Comentarios cualitativos
Rendimiento académico y valoración del mismo	Distribución del número y porcentaje de alumnos según la calificación. Proporción y rendimiento por cada vía planteada
Carga de trabajo para alumnado y profesorado	Profesorado: N° de horas por asignatura y mes. Alumnado: N° de horas por actividades y mes

Consta de un conjunto de cuatro apartados: (a) un grupo de ítems donde se caracteriza el contexto con las diferentes propuestas de evaluación que utiliza el profesor (tipo de asignatura, número de estudiantes, horas de clase, tutorías, actividad del alumnado, etc.) y donde se expone el tipo de evaluación a emplear, con las vías ofertadas y el número de estudiantes que se acogen a cada vía; (b) un grupo de ítems donde se expone la satisfacción con los procesos de evaluación seguidos. Se parte de un cuestionario anónimo de evaluación de la asignatura por parte del alumnado, con una escala likert de 5 niveles (0–4). En su validación se obtiene un Alfa de Cronbach elevada (0.84), con un índice KMO de 0.807 y en el test de esfericidad de Barlett un valor de $p > .00$. Lleva siendo utilizado tres cursos y se aplica de forma anónima a todo el alumnado en la última semana lectiva de la asignatura; (c) las calificaciones obtenidas por el alumnado, organizadas por totales y por parciales, según la vía de aprendizaje elegida. Los datos proceden de las actas de calificaciones oficiales de la asignatura; y (d) un grupo de ítems donde se expone la carga de trabajo correspondiente al profesorado y al alumnado.

Procedimiento y análisis

Cada profesor se compromete a desarrollar un sistema de evaluación formativo en, al menos, una asignatura que imparta durante el curso, ya sea en el primer o en el segundo cuatrimestre. Al principio del periodo lectivo cada profesor recoge la información relativa al diseño de la asignatura, que corresponde con el apartado (a) del instrumento. A lo largo de todo el cuatrimestre, tanto el alumnado como el profesorado que imparte la asignatura recogen de forma sistemática el número de horas que va dedicando a las actividades de enseñanza- aprendizaje y evaluación, a través de instrumentos de anotación y cálculo de carga de trabajo, diseñados ‘ad hoc’ y que se complementan y siguen semanalmente. Al final del cuatrimestre se completa el instrumento con los resultados académicos obtenidos y con los datos recogidos a través de su aplicación.

Una vez completado el informe estructurado del profesor se vuelcan y analizan a través del programa SPSS 18.0. Se aplican estadísticos descriptivos y ANOVAS. Los resultados se ponen en común por los profesores que forman parte del proyecto de *I + D*, con el fin de mejorar los sistemas de evaluación utilizados, particularmente con la viabilidad de la EF y con la forma de solucionar los problemas encontrados en su aplicación.

Resultados y discusión

Aprendizaje y satisfacción alumnado

Hemos encontrado valoraciones muy altas del alumnado en los ítems que hacen referencia al aprendizaje realizado a través de procesos de EF, así como la satisfacción con los mismos. En la [Tabla 4](#) presentamos las puntuaciones medias de las valoraciones otorgadas por el alumnado sobre algunos aspectos de los

Tabla 4. Media de las valoraciones otorgadas por el alumnado sobre la evaluación utilizada por el profesorado (escala likert de 0 a 4).

Términos utilizados para valorar los aprendizajes	Media
5.- El estudiante realiza un aprendizaje activo	3.42
6.- El alumno está más motivado, el proceso de aprendizaje es más motivador	3.08
7.- La calificación es más justa	3.13
8.- Mejora la tutela académica (seguimiento y ayuda al alumno)	3.03
9.- Permite aprendizajes funcionales	3.08
10.- Genera aprendizajes significativos	3.41
11.- Se aprende mucho más	3.32

procesos de evaluación formativa realizados por el profesorado, utilizando una escala Likert en la que 4 es el valor máximo.

Se agrupan estos ítems en dos conjuntos: los referidos al aprendizaje (5, 9, 10 y 11) y a la satisfacción con los procesos de EF seguidos (ítems 6, 7 y 8). Las valoraciones del alumnado sobre el aprendizaje percibido se pueden considerar muy altas, especialmente elevadas en aspectos como ‘la realización de un aprendizaje activo’ (3.42), la generación de aprendizajes significativos (3.41), y elevado grado de acuerdo con que ‘se aprende mucho más’ que con modelos más tradicionales (3.32). En cuanto a las valoraciones sobre su satisfacción con este tipo de procesos de aprendizaje y evaluación, sus valores también son muy altos, destacando el que señala una mayor percepción de justicia en la calificación de la asignatura. Estos resultados coinciden con los encontrados por Dunn y Mulvenon (2009), López-Pastor *et al.* (2013) y Taras (2005), que documentan la mejora de los resultados educativos a través de la EF. Por otra parte, Álvarez, Grau, y Tortosa (2010), tras el desarrollo de sistemas de EF en la universidad, concluyen que se ha incrementado la asistencia a clase, la implicación de los alumnos en las tareas, se ha mejorado la calidad de las tareas y trabajos realizados, así como los resultados académicos. También para Hernández (2012) es más eficaz y eficiente la evaluación que da retroalimentación al alumno para facilitar su aprendizaje.

Rendimiento académico

Según lo mostrado en la [Tabla 5](#), el rendimiento general podría valorarse como bueno, siendo el porcentaje de alumnado que supera la asignatura bastante alto (en torno al 85%). El porcentaje de no presentados (7,8%) y de los presentados (7,6%)

Tabla 5. Rendimiento general en todos los casos, diferenciado por tipo de calificación.

	Calificaciones					
	NP	Suspense	Aprobado	Notable	Sobresaliente	Mat. Honor
Porcentajes	7.8	7.6	23	47.9	11.8	3.66

Tabla 6. Índice de rendimiento académico en función de la vía de aprendizaje y evaluación elegida.

Vía	Continua	Mixta	Examen	Total
Índice de rendimiento	1.91	1.48	.53	1.57
VALORES DEL ANOVA INTER-VÍAS				
Suma de cuadrados 29.98	<i>gl</i> 2	Media cuadrática 7.419	<i>F</i> 29.346	Sig. .000

en ambos es bajo. El mayor porcentaje de alumnado se encuentra en la calificación de notable, seguido de aprobado y sobresaliente. En las universidades españolas la normativa suele establecer que las ‘matrículas de honor’ sólo pueden concederse, como máximo, al 5% del alumnado en una asignatura.

Con estos datos hemos calculado un ‘índice de rendimiento académico’, aplicando la habitual fórmula de la universidad para los expedientes académicos del alumnado, consistente en calcular el sumatorio de multiplicar las calificaciones por una escala 0–4 (los no presentados y suspensos por 0, aprobados por 1, notables por 2, sobresalientes por 3 y matrículas de honor por 4) y dicho sumatorio dividirlo por el número de sujetos, de modo que se obtenga un índice entre 0 y 4. De esta forma, además de tener un índice de rendimiento global y por vías, hemos podido realizar ANOVAS para comprobar si las diferencias de rendimiento entre vías de evaluación son significativas. En la [Tabla 6](#) presentamos los índices de rendimiento en función de la vía de evaluación elegida por el alumnado, así como el resultado del ANOVA inter-vías.

La tabla nos permite realizar dos tipos de análisis: (a) hay una fuerte diferencia de rendimiento académico entre las vías, que es estadísticamente significativa; y (b) la vía que mayor rendimiento tiene es la vía continua, con una media cercana al notable, seguida de la vía mixta, con un valor ligeramente superior al aprobado. La vía que muestra un rendimiento académico más bajo es la vía de examen final, con valores claramente cercanos al suspenso.

Para analizar con mayor profundidad esta diferencia entre vías de aprendizaje y evaluación, en la [Tabla 7](#) presentamos la distribución de calificaciones del alumnado en función de la vía elegida.

Puede observarse que la vía examen final agrupa el mayor porcentaje de suspensos y no presentados, así como calificaciones muy bajas entre los que sí logran superar la asignatura por esta vía. Por otro lado, la mayor parte del alumnado que opta por la vía continua logra calificaciones elevadas (notable, sobresaliente o matrícula de honor) y el porcentaje de ‘no aptos’ es muy bajo, tanto en no presentados como en suspensos. La vía mixta tiene un porcentaje bastante más elevado de aptos que de ‘no aptos’ (29% frente a 6%), pero logra porcentajes bastante más bajos en las calificaciones elevadas; curiosamente las diferencias más fuertes con la vía continua se dan en el notable, mientras que los porcentajes son muy similares en el sobresaliente y cercanos en la matrícula de

Tabla 7. Porcentajes de calificaciones obtenidas por el alumnado en cada vía.

Vías	Calificación					
	N.P	Suspenso	Aprobado	Notable	Sobresaliente	Mat. Honor
Continua	2.6	2.1	10.8	28.7	7.0	2.54
Mixta	2.1	4.0	9.9	13.4	6.0	1.0
Examen final	5.0	4.9	10.11	1.1	–	–

honor. Estos resultados coinciden con los encontrados por Arribas (2012), López-Pastor (2008) y Zaragoza *et al.* (2008) sobre el buen rendimiento académico que suele obtenerse cuando se desarrollan sistemas de EF en docencia universitaria. Así mismo, los estudios de Buscá *et al.* (2010), Castejón *et al.* (2011), López-Pastor (2008, 2009), López-Pastor *et al.*, (2011, 2013) y Santos *et al.* (2009) también encuentran diferencias de rendimiento entre las vías de aprendizaje y evaluación elegidas por el alumnado, en el sentido de que el alumnado que sigue la vía de EF obtiene un mayor rendimiento académico que el alumnado que sigue la vía de examen final.

Carga de trabajo del profesorado

En la *Tabla 8* se exponen los resultados sobre la actividad del profesorado que ha utilizado la evaluación formativa.

Se observa una elevada media de horas de ‘preparación de las clases’ (PC), así como las dedicadas a la ‘corrección de documentación del alumnado’ (CD). Las medias más bajas son las dedicadas al ‘uso y mantenimiento de los campus virtuales’ (CV) (Moodle, WebCt, etc.) en el desarrollo de las materias. Tanto las horas totales como la distribución interna de CD muestran ciertas variaciones entre cuatrimestres.

La carga media de trabajo es de 122 horas para el primer cuatrimestre y 158 para el segundo. Si dividimos esto por las 15 semanas que tiene cada cuatrimestre, el resultado medio está entre las ocho y 10 horas semanales, lo cual no parece suponer una carga excesiva de trabajo. Hay que tener en cuenta que se trata de medias y que las variaciones entre los casos pueden ser notables. La distribución

Tabla 8. Carga de trabajo promedio en horas para el profesorado repartidas para el total del estudio y en cada cuatrimestre.

Tipo de actividad	Totales 1er. cuatrimestre	Totales 2º cuatrimestre
PC	50.9	48.7
CD	47.1	84.5
CV	24	24.6
Totales	122	158

(PC = preparación de clases; CD = corrección de documentos; CV = Campus virtual)

de carga de trabajo se acumula sobre todo en los meses finales de cada cuatrimestre. Hernández (2012) observa la preocupación de los docentes por el alto grado de trabajo que implica la implementación de EF de carácter continuo, debido a la necesidad de proporcionar a los estudiantes retroalimentación efectiva en el momento oportuno y al gran número de alumnos implicados, en comparación con sistemas tradicionales.

Carga de trabajo del alumnado

La media de carga total de trabajo del alumnado es de aproximadamente 85 horas anuales, por tanto, en torno a 42.5 horas por cuatrimestre. La mayor parte de las asignaturas cuatrimestrales son de seis créditos y con el actual sistema ECTS cada crédito equivale a una media de 25 horas de trabajo del alumnado, de las cuales 10 suelen ser lectivas y 15 de trabajo autónomo y estudio por parte del estudiante. Por tanto, una asignatura de seis créditos supone un total de 90 horas de trabajo autónomo del alumnado. Dado que la carga de trabajo media es de 42.5 horas, se podría considerar que el desarrollo de sistemas de EF en la docencia universitaria no supone una sobrecarga de trabajo para el alumnado, sino que está claramente dentro de lo establecido en el sistema ECTS.

También hemos analizado la distribución de la carga de trabajo del alumnado, organizada en función de las actividades de aprendizaje de la asignatura (ver [Tabla 9](#)). Los datos se presentan en porcentajes medios respecto al tiempo total utilizado.

El porcentaje más alto corresponde a la realización de trabajos e informes de los temas (un tercio de la carga total). A continuación se encuentran las lecturas y recensiones y la preparación y aplicación de sesiones; dejando en último lugar la carga correspondiente a las revisiones y reconducciones de los trabajos.

Esta misma carga de trabajo distribuida por meses ([Tabla 10](#)) nos lleva a afirmar que la mayor carga de trabajo para el alumnado ocurre siempre en el último mes del cuatrimestre, mientras que la menor carga se da en el primer mes; sin embargo, en los meses intermedios se distribuye de forma bastante equilibrada.

De acuerdo con esos resultados, López-Pastor *et al.* (2013) observan disparidad entre las opiniones subjetivas de los estudiantes y profesores con respecto a los datos reales de la carga de trabajo recogidos de forma sistemática, dentro de los patrones establecidos oficialmente. Para estos autores, los sistemas de EF

Tabla 9. Carga de trabajo para el alumnado (con un promedio de 87 horas anuales/signatura) según la actividad propuesta por el profesor.

Tiempo de trabajo para:	Total
Lecturas y recensiones	24.1%
Preparación y aplicación de sesiones prácticas	23.0%
Trabajos e informes de temas	33.3%
Revisiones y reconducciones de los trabajos (grupal y en tutoría)	19.5%

Tabla 10. Carga de trabajo para el alumnado, distribuida por meses y cuatrimestres (en horas).

	sep	Oct	nov	dic	ene	feb	mar	abr	may
Tiempo de trabajo	10.1	19.5	19.2	22.7	34.1	10.8	13.5	12.8	22.3

suponen una carga de trabajo mucho mayor que los sistemas tradicionales de examen final, donde el alumno solía limitarse a estudiar el último mes y el profesor a corregirle. Por tanto, la percepción del profesorado y el alumnado se explica más por la comparación que realizan respecto al modelo tradicional, que por las evidencias encontradas sobre la carga de trabajo real que supone. Unos datos similares son encontrados por Julián *et al.* (2010), al indicar que la aplicación de la EF en docencia universitaria es perfectamente viable dentro de las cargas de trabajo que establece el sistema ECTS para el alumnado y el profesorado.

Conclusiones

El objeto del estudio es analizar la influencia que tiene la aplicación de sistemas de EF en la docencia universitaria en la satisfacción del alumnado con los mismos, en la mejora del rendimiento académico del alumnado y en la carga de trabajo de alumnado y profesorado.

Los resultados muestran valoraciones muy altas del alumnado en los ítems que hacen referencia al aprendizaje realizado a través de procesos de EF, así como la satisfacción con los mismos.

En lo relativo al rendimiento académico, aparecen diferencias estadísticamente significativas entre las vías de aprendizaje y evaluación, siendo la vía de EF y continua la que obtiene un mayor rendimiento, seguida de la vía mixta. La vía de examen final muestra un rendimiento muy bajo.

En cuanto a la carga de trabajo del profesorado, el resultado medio está entre las ocho y 10 horas semanales de trabajo docente no lectivo (preparación de clases, corrección de documentos y actividades del alumnado, mantenimiento de la web y tutoría virtual con el alumnado). La actividad que menos tiempo parece requerir es la tercera.

Por último, la carga de trabajo media del alumnado es de 42.5 horas por asignatura cuatrimestral (seis créditos), que está claramente dentro de lo establecido en el sistema ECTS, por lo que se podría considerar que el desarrollo de sistemas de EF en la docencia universitaria no supone una sobrecarga de trabajo para el alumnado.

En cuanto a las limitaciones del estudio, hemos comprobado que al permitir que el alumnado elija las diferentes vías de evaluación, la mayoría de los buenos alumnos suelen elegir la vía continua, por lo que dicha variable puede estar afectando al resultado. Se trata de un problema de difícil solución debido a la

dificultad de implementar otro tipo de diseños (experimentales y cuasi-experimentales). Una limitación de orden menor se refiere al tamaño de la muestra, en el sentido de que sería interesante acumular datos con poblaciones más amplias.

Estos resultados muestran la conveniencia de ir avanzando en el desarrollo de sistemas de EF en la docencia universitaria, por las considerables mejoras que supone en el rendimiento académico, el aprendizaje y la satisfacción del alumnado. Por otra parte, no parecen suponer una carga de trabajo tan excesiva para alumnado y profesorado, como suele plantearse desde sus creencias. En consecuencia, se sugiere la aplicación de instrumentos y procesos de evaluación formativa de manera extensiva en la docencia universitaria, de modo que no se trate de un sistema de evaluación opcional en las guías docentes, sino una propuesta potenciada y avalada desde las instituciones académicas.

Acknowledgements / Agradecimientos

This article is part of the EDU 2010–19637 Research Project (EDUC Sub-programme), titled: *La evaluación en la formación inicial del profesorado de educación física (Assessment in initial teacher training of physical education teachers)*. National Plan for R&D&I Projects (2008–2011). / Este artículo forma parte del Proyecto de Investigación EDU 2010–19637 (Subprograma EDUC), titulado: *La evaluación en la formación inicial del profesorado de educación física. Plan nacional de Proyectos I + D + i (2008–2011)*.

References / Referencias

- Álvarez, J. D., Grau, S., & Tortosa, M. T. (2010). Estrategias de coordinación metodológicas en la evaluación formativa de una asignatura. In C. Gómez & S. Grau (Eds.), *Evaluación de los aprendizajes en el EEES* (pp. 75–90). Alcoy: Marfil.
- Arribas, J. M. (2012). El rendimiento académico en función del sistema de evaluación empleado. *RELIEVE*, 18–1, art. 3. Retrieved from http://www.uv.es/RELIEVE/v18n1/RELIEVEv18n1_3.htm
- Bennet, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18, 5–25. doi:10.1080/0969594X.2010.513678
- Black, P., Harrison, C., Hogden, J., Marshall, B., & Wiliam, D. (2005). Dissemination and evaluation: A response to Smith and Gorard. *Research Intelligence*, 93, 6–7.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5, 7–74. doi:10.1080/0969595980050102
- Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education. Learning for the long term*. London: Routledge.
- Brown, S., & Glasner, A. (2003). *Evaluar en la universidad. Problemas y nuevos enfoques*. Madrid: Narcea.
- Buscá, F., Pintor, P., Martínez, L., & Peire, T. (2010). Sistemas y procedimientos de Evaluación Formativa en docencia universitaria: Resultados de 34 casos aplicados durante el curso académico 2007–2008. *Estudios Sobre Educación*, 18, 255–276.
- Castejón, F. J., López-Pastor, V. M., Julián, J., & Zaragoza, J. (2011). Evaluación formativa y rendimiento académico en la Formación inicial del profesorado de Educación Física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 11, 328–346.

- Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7), 1–11.
- Falchikov, N. (2005). *Improving assessment through student involvement. Practical solutions for aiding learning in higher and further education*. London: Routledge.
- Fernández-Pérez, M. (1989). *Así enseña nuestra universidad*. Salamanca: Hispagraphis.
- Fraile, A. (2010). La autoevaluación: Una estrategia docente para el cambio educativo de valores en el aula. *Ser Corporal, 3*, 6–18.
- Fraile, A. (2012). Evaluación formativa e interdisciplinarietà: Análisis de dos asignaturas con el mismo sistema de evaluación. *Psychology, Society, & Education, 3*(2), 5–16.
- Gibbs, G., & Simpson, C. (2004–05). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3–31.
- Gibbs, G. (2003). Uso estratégico de la evaluación en el aprendizaje. In S. Brown & A. Glasner (Eds.), *Evaluar en la universidad. Problemas y nuevos enfoques* (pp. 61–75). Madrid: Narcea.
- Greenwald, A. G., & Gillmore, G. M. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology, 89*, 743–751. doi:10.1037/0022-0663.89.4.743
- Gutiérrez, C., Pérez-Pueyo, A., Pérez- Gutiérrez, M., & Palacios, A. (2011). Percepciones de profesores y alumnos sobre la enseñanza, evaluación y desarrollo de competencias en estudios universitarios de formación del profesorado. *Cultura y Educación, 23*, 499–514.
- Hernández, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education, 64*, 489–502. doi:10.1007/s10734-012-9506-7
- Julián, J., Zaragoza, J., Castejón, F. J., & López-Pastor, V. M. (2010). Carga de trabajo en diferentes asignaturas que experimentan el sistema ECTS. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte, 10*, 218–233.
- López-Pastor, V. M. (2008). Desarrollando sistemas de evaluación formativa y compartida en la docencia universitaria. Análisis de resultados de su puesta en práctica en la formación inicial del profesorado. *European Journal of Teacher Education, 31*, 293–311. doi:10.1080/02619760802208452
- López-Pastor, V. M. (Ed.) (2009). *Evaluación Formativa y Compartida en Docencia Universitaria*. Madrid: Narcea.
- López-Pastor, V. M., Manrique, J. C., & Vallés, C. (2011). Evaluación y calificación en los nuevos estudios de Grado. Especial incidencia en la Formación Inicial del Profesorado. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 39*(14.4), 1–20. Retrieved from <http://www.aufop.com/aufop/revistas/arta/digital/165/1675>
- López-Pastor, V., Pintor, P., Muros, B., & Webb, G. (2013). Formative assessment strategies and their effect on student performance and on student and tutor workload: The results of research projects undertaken in preparation for greater convergence of universities in Spain within the European Higher Education Area (EHEA). *Journal of Further and Higher Education, 37*, 163–180. Retrieved from <http://dx.doi.org/10.1080/0309877X.2011.644780>
- Manrique, J. C., Vallés, C., & Gea, J. M. (2012). Resultados generales de la puesta en práctica de 29 casos sobre el desarrollo de sistemas de Evaluación formativa en docencia universitaria. *Psychology, Society & Education, 4*(1), 83–98.
- MEC. (2006). *Propuestas para la renovación de las Metodologías Educativas en la Universidad*. Madrid: MEC.
- Navarro, V., Santos, M. L., Buscá, F., Martínez, L., & Martínez, L. F. (2010). La experiencia de la Red Universitaria Española de Evaluación Formativa y

- Compartida: Proceso y abordaje. *Revista Iberoamericana de Educación*, 52(7), 1–12. Retrieved from <http://www.rieoei.org/expe/3428Navarro.pdf>
- Palacios, A., & López-Pastor, V. M. (2013). Haz lo que yo digo pero no lo que yo hago: Sistemas de evaluación del alumnado en la formación inicial del profesorado. *Revista de Educación*, 361, 279–305. doi:10.4438/1988-592X-RE-2011-361-143
- Rué, J. (2009). *El aprendizaje autónomo en la Educación Superior*. Madrid: Narcea.
- Santos, M. L. Martínez, L. F., & López-Pastor, V. M. (Eds.) (2009). *La innovación docente en el EEES. Experiencias de evaluación formativa y compartida en la formación inicial del profesorado*. Almería: Universidad de Almería.
- Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: Guilford Press.
- Taras, M. (2005). Assessment, summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53, 466–478. doi:10.1111/j.1467-8527.2005.00307.x
- Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of Further and Higher Education*, 33, 57–69. doi:10.1080/03098770802638671
- Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education*, 38, 323–342. doi:10.1080/03054985.2012.689693
- Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27, 615–631. doi:10.1080/01411920120095780
- Trillo, F., & Porto, M. (1999). La percepción de los estudiantes sobre su evaluación en la universidad. Un estudio en la Facultad de Ciencias de la Educación. *Innovación Educativa*, 9, 55–75.
- Winchester, M. K., & Winchester, T. M. (2012). If you build it will they come? Exploring the student perspective of weekly student evaluations of teaching. *Assessment & Evaluation in Higher Education*, 37(6), 671–682. doi:10.1080/02602938.2011.563278
- Zabalza, M. A. (2001). Evaluación de los aprendizajes en la Universidad. In A. García-Valcárcel (Ed.), *Didáctica Universitaria* (pp. 261–291). Madrid: La Muralla.
- Zaragoza, J., Luis-Pascual, J. C., & Manrique, J. C. (2008). Experiencias de innovación en docencia universitaria: Resultados de la aplicación de sistemas de evaluación formativa. *REDU. Revista De Docencia Universitaria*, 7(4), 1–33. Retrieved from <http://redaberta.usc.es/redu/index.php/REDU/article/view/111>