ongitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

Estudio longitudinal sobre los efectos del desarrollo de la Evaluación Formativa y Compartida en la Formación Inicial del Profesorado

Cristina Pascual-Arias¹ (b)
Víctor M. López-Pastor¹ (b)
David Hortigüela-Alcalá² (b)

¹ Education Faculty, University of Valladolid, Segovia, Spain ² University of Burgos, Burgos, Spain

Correspondence: Cristina Pascual-Arias,

Cristina Pascual-Arias, cristina.pascual@uva.es

Short title:

Longitudinal effects of Formative And Shared Assessment

How to cite this article:

Pascual-Arias, C., López-Pastor, V. M., & Hortigüela-Alcalá, D. (2023). Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education. *Cultura, Ciencia y Deporte, 18*(55), 5-31. https://doi.org/10.12800/ccd.v18i55.1904

Received: 30 March 2022 / Accepted: 12 September 2022

Abstract

This study presents the research outcomes of a Formative and Shared Assessment experience in future Early Childhood Education teachers over two consecutive academic years, paying special attention to the assessment system's advantages and disadvantages, student's academic performance and teacher and student workload. It was a developmental longitudinal ex-post-facto study conducted in one cohort of pre-service teacher education (PSTE), applied to two consecutive courses of the Early Childhood Education Degree related to Physical Education. The data collection instruments were an anonymous student questionnaire, validated by Castejón-Oliva et al. (2015) and the best practices reports proposed by the Formative and Shared Assessment Network, filled in by the course teachers. The results revealed that students found more advantages than disadvantages in the assessment system in both courses. Besides, academic performance was observed to be very positive, and it was higher in the second course, compared to the first one. Furthermore, there was a high correlation between workload and academic performance in both courses.

Keywords: formative assessment, shared assessment, pre-service teacher education, academic performance.

Resumen

Este estudio presenta los resultados de investigación de una experiencia de Evaluación Formativa y Compartida en los futuros maestros de Educación Infantil durante dos cursos consecutivos; prestando especial atención a las ventajas e inconvenientes del sistema de evaluación, rendimiento académico del alumnado y carga de trabajo del profesorado y del alumnado. Es un estudio ex-post-facto evolutivo longitudinal en una misma cohorte de Formación Inicial del Profesorado (FIP), del Grado de Educación Infantil, vivenciado en dos asignaturas sucesivas del ámbito de la Educación Física. Los instrumentos de recogida de datos que se han utilizado son un cuestionario anónimo del alumnado, validado por Castejón-Oliva et al. (2015) y los informes de buenas prácticas de la Red de Evaluación Formativa y Compartida realizados por los docentes que impartieron estas asignaturas. Los resultados muestran cómo el alumnado encuentra más ventajas que inconvenientes en el sistema de evaluación en ambas asignaturas. También se refleja que el rendimiento académico del alumnado es muy positivo, aumentando en el segundo curso con respecto al primero. Además, se ha obtenido una alta correlación entre la carga de trabajo y el rendimiento académico en ambas asignaturas.

Palabras clave: evaluación formativa, evaluación compartida, formación inicial del profesorado, rendimiento académico.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

6

Introduction

What is Formative and Shared Assessment?

The unification process that took place in the European Higher Education Area (EHEA) caused a shift to a 'Dialogical Learning' model (Aubert et al., 2008; López-Pastor, 2009). A thorough review of university education practices was conducted, among which the shift towards continuous assessment must be highlighted (Mérida, 2006; Santos et al., 2018).

Dochy et al. (2002) established that universities needed to move from a 'culture of exams' to a 'culture of assessment' by means of continuous assessment. In this regard, the proposal called Formative and Shared Assessment (FSA) has generated a large number of best practices experiences (Gallardo-Fuentes & Carter-Thuillier, 2016; López-Pastor, 2009; Romero et al., 2014).

Several authors consider that FSA is the assessment methodology that best suits the new EHEA demands (Hortigüela-Alcalá et al., 2016; López-Pastor, 2009). López-Pastor (2009) defined Formative Assessment as any assessment process whose aim is to improve the teachinglearning process through the following three aspects: (a) the students' learning; (b) the teaching practice; (c) the teaching-learning process, by fixing the errors detected. The concept of Shared Assessment refers to the student's involvement in the assessment process. This teacherstudent dialogue can be individual or in groups and is usually linked to previous self-assessment and/or peerassessment processes. This concept is only related to formative assessment, i.e. learning-oriented assessment. When referring to marking, concepts such as self-marking (by the student), dialogued marking (teacher-student) and peer marking (between students) should be used (López-Pastor, 2009).

Formative and Shared Assessment in Pre-Service Teacher Education

Various studies support the benefits of implementing FSA in pre-service teaching education (PSTE): (a) it increases the student's motivation towards and engagement in their learning (Gallardo et al., 2018; López-Pastor et al., 2020); (b) it develops the student's self-criticism, autonomy, responsibility and critical analysis (López-Pastor, 2009); (c) it helps the student overcome their challenges during the teaching-learning process (Barrientos et al., 2019; Córdoba et al., 2016); (d) it improves aspects like learning, academic performance and competency acquisition (Fraile-Aranda et al., 2013; Hortigüela-Alcalá et al., 2015a; Romero-Martín et al., 2014); (f) it is the way of learning that is most in line with EHEA dialogical learning (Hortigüela-Alcalá et al., 2019). Those teachers who apply the Formative Assessment in PSTE seem to be highly satisfied with this assessment system, despite it sometimes meaning a higher workload for them (Gutiérrez-García et al., 2011; Romero-Martín et al., 2015).

When looking at Early Childhood PSTE (PSTE-EC), Mérida (2006) pointed out that university education and working life may be disconnected. PSTE-EC focuses on providing students with a wide range of knowledge, but it prevents them from acquiring the knowledge that will allow them to solve situations in their professional life. This may be because, in PSTE, contents are sometimes taught without context and are far from the actual professional practice. In this regard, it would be advisable that FSA implementation in PSTE-EC provided the necessary teaching skills so that students could extrapolate this knowledge to their future professional practice (Basera, 2019).

Martínez-Mínguez and Flores (2014) conducted a study on PSTE-EC teaching guides and reported that it is necessary for students to get to know the assessment in a timely, clear and transparent manner for the teacher to be able to guide the theory and practice in a critical and thoughtful way.

How Formative and Shared Assessment Works in Pre-Service Teacher Education

Silva and López-Pastor (2015) gathered PSTE-EC students' perspectives on the FSA experienced during the degree and reported several positive aspects: (a) they deemed the effects of FSA on PSTE-EC as positive; (b) they stated that the FSA system helped them acquire teaching competencies. They also highlighted some drawbacks: (a) they were sometimes confused by FSA; (b) they showed some fears, resistance and doubts towards FSA; (c) they encountered difficulties in self- and peer-assessment. They considered that FSA would be more feasible if it was applied in all courses of the same university semester or degree.

Gallardo et al. (2018) conducted a study on the application of the FSA system in a PSTE-EC course, where the students gave high scores for the system's advantages and low scores for most of the disadvantages. Hortigüela-Alcalá et al., (2015a) researched the effects of applying FSA to different courses, but not with the same students. The students deemed the feedback received and the academic performance achieved through the FSA system as positive, compared to the traditional one. In this regard, the student's engagement in their own assessment and their reflection is essential (Southcot & Crawford, 2018). Gallardo et al. (2018) and Hortigüela-Alcalá et al. (2015b) believe that future teachers will apply with their students the FSA systems they experienced during PSTE. An example is the study by Pascual-Arias et al. (2019), based on the application of FSA systems in Early Childhood after having experienced them in their PSTE-EC. It yielded positive effects on the learning process quality, and children improved their expression and decision-making in the classroom.

In the beginning, the FSA system usually seems to be complex for students but later they deem some of its

advantages as positive, especially the feedback they receive from the teacher, their involvement in the assessment process and the academic performance achieved (Hortigüela-Alcalá et al., 2019; López-Pastor & Sicilia-Camacho 2017; Martínez-Mínguez et al., 2019; Romero-Martín et al., 2015). Hortigüela-Alcalá et al., (2015b) stated that students who had previously received FSA perceived it more favourably. Previous studies have proved that FSA considerably improves PSTE students' academic performance, compared to more traditional systems (Fraile-Aranda et al., 2013; Hortigüela-Alcalá et al., 2015a; López-Pastor et al., 2013; Panadero et al., 2014; Romero-Martín et al., 2014).

In the review conducted, no study was found to have applied FSA systems to the same cohort of PSTE-EC students. Therefore, it seemed interesting to examine the effects of FSA over time, especially on variables related to the teaching-learning process, students' academic performance and teacher workload.

To do so, the following research questions were posed:

RQ1: What happens when the same FSA system is applied to the same cohort of Early Childhood Teacher Education Degree students during two consecutive academic years, in two different but supplementary courses?

RQ2: What are the advantages and disadvantages? RQ3: How does FSA affect students' academic performance and teacher and student workload?

The initial hypothesis was that these variables may be positively affected by the previous experience gained with the same assessment system. That is, in the second year, the advantages would be more positively and the disadvantages would be less negatively perceived, academic performance would improve and (student) workload would decrease.

Method

An empirical study was conducted applying a quantitative methodology; it was a developmental longitudinal ex-post-facto study (Montero & León, 2007). This design allowed us to monitor the same group during two consecutive academic years, in order to compare a set of variables and their evolution over time.

Approval was obtained from the ethics committee CEICA (Research Ethics Committee of the Community of Aragon) C.P.-C.I.PI21/377.

Sample

The study participants belonged to the same cohort of the Early Childhood Teacher Education Degree, but students from two different paths came together in the second year. There were 41 students in the first year (3rd year of Early Childhood Education Degree) and 52 in the second (4th year of Early Childhood Education Degree and of the combined degree: Early Childhood and Primary Education Degree) (see Table 1).

Table 1. Characteristics of the courses analysed in the developmental longitudinal study

Course	Year and semester	ECTS	In-person hours	Students	Sex	Age
Fundamentals and Teaching of Early Childhood Physical Education (FTECPE)	3rd year of Early Childhood Education Degree, 2nd semester	6	Theory: 15 Practice: 30 Seminar: 15	41	Male: 5.3% Female: 94.7%	<20 years: 2.6% 20-25 years: 81.6% >25 years: 15.8%
Body Expression and Communication in Early Childhood Education (BECEDE)	4th year of Early Childhood Education Degree, 1st semester	6	Theory: 15 Practice: 30 Seminar: 15	52	Male: 10.6% Female: 89.4%	20-25 years: 89.4% >25 years: 10.6%

Source: own elaboration.

The students of the combined degree had completed the same course and used the same assessment system three years earlier.

Therefore, despite following a different path, they experienced the same course and assessment system with the same teachers. This organisation was based on the structure of the centre under study.

Table 2. Types of learning, assessment and marking offered within the FSA system

	Formative and continuous	Mixed	Final
Requirements	 Mandatory attendance. Only 15% of justified absence. To complete all learning activities correctly and in a timely manner. 	- Minimum attendance of 50%. - SLP is mandatory; the rest of activities are optional.	- Less than 50% of attendance SLP is mandatory.
Learning activities and contribution to the mark	- Supervised Learning Project (SLP) (35% of the mark) Practical session sheets + conceptual maps + reading reviews + monographic projects (45% of the mark).	- SLP (30% of the mark). - Same, but they are optional (20% of the mark).	- The SLP can be made during the course or it can be the third part of the final test (30% of the mark). - Practical test (20% of the mark).
Tests	- Theoretical tests with immediate peer assessment (20% of the mark).	- Theoretical test (50% of the mark).	- Theoretical test (50% of the mark).

Source: own elaboration.

The course Fundamentals and Teaching of Early Childhood Physical Education (FTECPE) focused on building professional skills related to body and psychomotor development, as well as movement teaching in Early Childhood education. The course Body Expression and Communication in Early Childhood Education (BECEDE) followed the previous one and focused on the knowledge and development of body language and motor game educational possibilities in Early Childhood education. The courses were consecutive and supplementary, and very similar learning and assessment activities were used.

The same FSA system was applied in both courses, and students were offered three types of learning and assessment: (a) formative and continuous; (b) mixed; or (c) final. The FSA process was unique and common to the three groups. However, the requirements to belong to every group, as well as the mark composition (according to the learning type followed) were different (see Table 2). Similar FSA systems can be found in Gallardo-Fuentes and Carter-Thuillier (2016) and López-Pastor (2009).

The teacher's role during the FSA process was to guide students through their learning process and to provide them with feedback for continuous improvement. Teaching was organised as follows: (a) the first course was given by one single teacher with broad experience in the content and in FSA; (b) the second course was given by the same teacher plus two other pre-service teachers with less experience (1 and 3 years, respectively). The teachers worked in a coordinated manner and there was one single teaching-learning and assessment process.

Data Collection Instruments

Two instruments were used for data collection:

Anonymous questionnaire on PSTE methodology and assessment. Previously validated (Castejón-Oliva et al., 2015), with a Cronbach's alpha reliability coefficient of .84 for the complete scale. Thirty items were used from this anonymous questionnaire, which were scored on

a five-point Likert-type scale (1=not at all; 2=slightly; 3=moderately; 4=very much; 5=extremely) in order to assess the FSA system's advantages (16 items) and disadvantages (14 items).

Best Teaching Practices Report (Hortigüela-Alcalá et al., 2019; Romero-Martín et al., 2014). They were semi-structured reports divided into two parts: (a) FSA experience description; (b) outcomes obtained (advantages, disadvantages and potential solutions, academic performance, workload and conclusions). The report of every course was used to collect academic performance and student and teacher workload data. It was filled in by the course teachers. There was only one report per course, regardless of the number of teachers.

Data Analysis Procedure

First of all, the sample distribution was analysed using the Kolmogorov-Smirnov test, since n>50. The academic performance and workload variables did not follow a normal distribution, so non-parametric tests were conducted for these variables. By contrast, the variables related to the FSA system's advantages and disadvantages did follow a normal distribution and parametric tests were applied in this case.

After this check, the data collected were analysed

through descriptive statistics (mean [M] and standard deviation [SD]) and inferential statistics (difference of means, Mann-Whitney U test, ANOVA and Pearson's correlation) using the statistical software SPSS 24.0.

The following categories were established for data analysis, according to the study aim: (a) FSA system's advantages (16 questionnaire items); (b) disadvantages (14 questionnaire items); (c) academic performance (4 items of the best practices report); (d) student and teacher workload (3 report items).

Results

The results have been organised based on the categories established and are described below.

Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

Perception of the Advantages of the FSA System Applied

Table 3 contains the results relative to the advantages of the FSA system applied in both courses.

All FSA system's advantages were very highly scored in both courses, although most of the items received slightly lower scores in the 4th-year course. These differences were only statistically significant in seven items: (a) it offers alternatives to all students; (b) it is process-centred, daily work is important; (c) students learn much more; (d) theory and practice are interrelated; (e) all possible aspects (as regards knowledge, know-how and know-how-to-be-and-behave) are assessed; (f) feedback is provided on documents and activities; (g) more responsibility is required.

Table 3. Results related to the FSA system's advantages (1-5 scale) (* indicates significant differences)

	FTE	CPE	BEC	EDE		
Questions	М	SD	М	SD	U	Sig.
1. It offers alternatives to all students	4.38	.854	4.07	.625	636.000	.006*
2. There is a previous, negotiated and agreed contract concerning the assessment system	4.33	.944	4.18	.870	804.500	.197
3. It is process-centred, daily work is important	4.93	.258	4.68	.518	731.000	.007*
4. The student learns actively	4.84	.374	4.84	.370	942.500	.963
5. Teamwork is presented from a collaborative perspective	4.72	.504	4.61	.618	857.000	.355
6. The student is more motivated, the learning process is more motivating	4.42	.626	4.43	.625	935.000	.917
7. The mark is fairer	4.26	.790	4.25	.719	922.000	.821
8. It enhances academic supervision (follow-up and support to the student)	4.35	.613	4.11	.618	761.000	.076
9. It allows for functional and significant learning	4.63	.489	4.39	.618	758.000	.068
10. Students learn much more	4.63	.618	4.41	.542	722.500	.029*
11. The quality of the requested projects increases	4.49	.631	4.27	.544	754.000	.054
12. Theory and practice are interrelated	4.72	.591	4.36	.685	627.000	.002*
13. All possible aspects (as regards knowledge, know-how and know-how-to-be-and-behave) are assessed	4.60	.627	4.23	.522	634.000	.004*
14. Feedback is provided on documents and activities	4.79	.412	4.30	.509	494.500	<.001*
15. A more individualised follow-up is provided	4.16	.721	3.89	.722	747.000	.062
16. More responsibility is required	4.74	.441	4.43	.545	666.500	.005*

Source: own elaboration.

Note: U=Mann-Whitney U; <math>M=Mean; SD=Standard deviation; *=<math>p<.05.

Perception of the Disadvantages of the FSA System Applied

Table 4 shows the data relative to the FSA system's disadvantages.

The results are much more scattered than in the advantages table; they can be divided into three groups in both courses: (a) 4 items with high scores in both courses (higher than 4); (b) 4 items in the first course and 1 item in the second one obtained average scores (3-4); (c) 6 items in the first course and 9 items in the second one obtained low scores (lower than 3).

The same trend was observed, i.e. the items received a lower score in the second course, although the differences were only statistically significant in five items: (a) the working method is not very well known, lack of habit; (b) it requires continuity; (c) it is more demanding; (d) there

is a mismatch between work and credits; (e) it is unfair, compared to other assessment processes. Statistically significant differences with high values were found in items 3, 5 and 8, meaning that students considered them as clear disadvantages, although the value was lower in the second course.

Items 2 and 11 also presented statistically significant differences, but the values were lower in both courses. This means that they were not considered as disadvantages in the second course, which seems to indicate that students became familiar with these FSA aspects. Item 1 obtained a high score in the first course, which significantly decreased in the second one. The two items with the sharpest significant decreases were 8 and 11. They received high or average values in the first course, but average or low values in the second one.

Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

Table 4. Results related to the FSA system's disadvantages (1-5 scale) (* indicates significant differences)

	FTECPE		ВЕ	CEDE		
Questions	М	SD	М	SD	U	Sig.
1. Attendance and active participation are mandatory	4.65	.613	4.64	.718	942.500	.970
2. The working method is not very well-known, lack of habit	3.58	1.451	2.91	1.291	686.500	.024*
3. It requires continuity	4.91	.294	4.59	.583	686.000	.003*
4. It needs to be previously understood	3.71	.995	3.49	.910	771.000	.220
5. It is more demanding	4.53	.735	4.00	.915	620.500	.003*
6. It is difficult to work in groups	2.79	1.337	2.49	1.077	827.500	.383
7. A lot of work may accumulate at the end	3.19	1.484	2.59	1.127	729.500	.059
8. There is a mismatch between work and credits	3.60	1.545	2.63	1.479	568.500	.004*
9. The marking process is more complex and, sometimes, not very clear	2.53	1.437	2.02	1.199	719.500	.093
10. It generates insecurity and uncertainty, doubts about what needs to be done	2.63	1.310	2.33	1.040	807.500	.296
11. It is unfair, compared to other assessment processes	2.56	1.695	1.67	.969	657.500	.014*
12. The corrections have not been very clear	2.44	1.181	2.07	1.246	743.500	.073
13. Work assessment is subjective	2.55	1.310	2.47	1.279	866.000	.736
14. It requires participating in my own assessment (self-assessment)	4.43	.859	4.23	.812	772.500	.210

Source: own elaboration.

Note: U=Mann-Whitney U; <math>M=Mean; SD=Standard deviation; *=<math>p<.05.

Perception of Academic Performance

Table 5 contains the academic performance results in both courses.

The data revealed that the majority of students chose continuous learning and assessment in both courses (82.9% and 88.4%), but a slight inversion was found between the

mixed and final types from one course to the next one (Table 5). Large performance differences were observed among assessment types: all no-shows and fails belonged to the mixed and final types, while the majority of students who had chosen the continuous type received Noteworthy or Outstanding. Besides, performance was clearly higher in the second course, but the differences were not significant according to the ANOVA (0.087).

Table 5. Comparison of students' academic performance in the two courses (percentages)

					N	Mark		
Course	Type of assessment	NS	Fail	Pass (50- 69/100)	Noteworthy (70-89/100)	Outstanding (90-100/100)	Distin- ction	Total/Type
	Continuous	2.4	-	4.9	56.1	14.6	4.9	82.9
FTECPE	Mixed	2.4	-	4.9	-	-		7.3
(3rd year)	Final	4.9	4.9	-	-	-		9.8
	Total	9.7	4.9	9.8	56.1	14.6	4.9	100
	Continuous	-	-	1.9	53.8	26.9	5.8	88.4
BECEDE	Mixed	-	-	5.9	1.9	-	-	7.8
(4th year)	Final	3.8	-	-	-	-	-	3.8
	Total	3.8	-	7.8	55.7	26.9	5.8	100
	X (Avera	ge mar	'k*)	SD	S	ig.	F	n²
FTECPE (3rd year)	2.	80		1.209		207	2.006	2.444
BECEDE (4th year)	3.	.19 .951		0.0	087	2.996	3.441	

Source: own elaboration.

^{*}Average mark on a 0-5 scale (0=no-show, 1=fail, 2=pass, 3=noteworthy, 4=outstanding, 5=distinction).

Perception of Teacher and Student Workload

Table 6 shows the teacher workload data. The results revealed that the number of hours spent by the teachers was considerably higher in the second course (164 hours) than in the first one (109 hours). Both had officially the same number of credits (6 ECTS credits) but, in the second

one, there was a slightly higher number of students and a larger number of teachers (1 vs. 3).

In both courses, the teaching tasks that required the highest number of hours were document correction and lesson preparation, taking twice as many hours in the second course as in the first one.

Table 6. Comparison of teacher workload in each course

	Course and hours of teacher's work per month											
Teaching tasks			FT	ECPE			,		BE	CEDE		
	PM	M1	M2	М3	M4	Total	PM	M1	M2	М3	M4	Total
Lesson preparation	10	4	4	4	4	26	15	10	10	10	10	55
Correcting	-	15	14	25	10	65	-	30	30	30	30	120
Web maintenance + e-mailing with students	2	4	4	4	4	18	6	1	1	1	1	10
Subtotals	12	23	23	33	18	109	25	41	41	41	41	164

Source: own elaboration.

Note: PM = previous months; M1 = month 1; M2 = month 2; M3 = month 3; M4 = month 4; T = total.

Table 7 presents a comparison of the number of hours of work spent by students. Student workload was very similar in both courses, but slightly higher in the first one.

Differences were detected in three types of activities: (a) readings and reviews (half load in the second course because fewer readings and reviews were completed in the second course); (b) projects and reports (almost twice as much load in the second course as in the first one because more

of these activities were completed); (c) project revision and correction (20 vs. 5 hours, many fewer hours in the second course, which seems to indicate that students delivered better-quality activities from the beginning, so that fewer corrections were needed afterwards). The mean number of hours of student work was larger in the 3rd-year course than in the 4th-year course, with large variability among students. The ANOVA revealed that the differences were significant (<.001), the workload being lower in the second course.

Table 7. Comparison of hours of work spent by students on each course

Student tasks		Course	e and hours of student	s work per n	nonth
	FTECPE		BECEDE		
Reading	s and reviews	20		10	,
Practical session p	reparation and reports	28		25	
Monographic p	projects and reports	18		30	
,	direction (group or individual essions)	20		5	
Study f	or the exam	15		15	
Su	Subtotals			85	
Course	X (Average workload)	SD	Sig.	F	n²
FTECPE (3rd year)	114.55	20.89	<.001	46.161	13606.811
BECEDE (4th year)	87.93	13.91	\. 001	40.101	13000.811

Source: own elaboration.

Table 8 shows a comparison between student workload and academic performance, to complete the analysis of both variables.

The mean workload was slightly higher in the 3rd-year course, although the extreme values (minimum and maximum) were similar in both of them. In both courses, a clear relationship

was observed between workload and academic performance; the better the mark, the higher the workload, except for Outstanding and Distinction in the first course. Furthermore, a large difference was detected between courses, as a much lower workload seemed to be necessary in the second course in order to obtain the same outcomes, especially Noteworthy and Outstanding (110 vs. 80 h and 150 vs. 100 h, respectively).

Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

Table 8. Comparison of the number of hours of work and students' academic performance

Items to be score	d regarding student workload	Course			
	FTECPE	BECEDE			
Average student indep	pendent workload apart from lesson hours	96 hours	85 hours		
Workload of the student	ss who did most and least in the group (in hours)	Minimum value:38 hours Maximum value: 210 hours	Minimum value: 39 hours Maximum value: 212 hours		
	Fail	38 hours	-		
	Pass (50-69/100)	60 hours	45 hours		
Mark and number of hours	Noteworthy (70-89/100)	110 hours	80 hours		
	Outstanding (90-100/100)	150 hours	100 hours		
	Distinction	115 hours	120 hours		

Source: own elaboration.

Table 9. Significance and correlation between workload and academic performance

Course	Pearson's correlation	Significance
FTECPE (3rd year)	0.730	<.001
BECEDE (4th year)	0.899	<.001

Source: own elaboration.

The correlation analysis (Table 9) yielded significant correlations in both courses (α <.001), especially strong in the second one (0.899).

Discussion

The aim of this study was to examine a cohort of PSTE-EC students in order to understand how an FSA system evolved and how it affected variables like advantages and disadvantages, academic performance and workload. Two different courses belonging to the same knowledge field and taught in consecutive semesters were analysed and compared.

The students from both courses reported many more advantages than disadvantages in the FSA process, as in the study by Gallardo et al. (2018). Nevertheless, in this study, significant differences were found in other advantages and disadvantages. The advantages obtained very high scores, most of them above 4 on a 1-5 scale. The major advantages were in line with those found in the studies by López et al. (2008) and Romero-Martín et al. (2014).

Statistically significant differences between courses were observed in seven advantages: (a) it offers alternatives to all students; (b) it is process-centred, daily work is important; (c) students learn much more; (d) theory and practice are interrelated; (e) all possible aspects (as regards knowledge, know-how and know-how-to-be-and-behave) are assessed; (f) feedback is provided on documents and activities; (g) more responsibility is required. The study by Atienza et al. (2016) revealed very high scores on the advantages analysed, such as motivation towards the course and

responsibility development. This could be, as stated by the authors, because FSA fosters active learning and, therefore, students feel more motivated, autonomous and responsible. Nonetheless, these authors only examined the scores in one course.

In the present study, all scores related to the advantages were lower in the second course. This is in contrast with the studies by Gallardo et al. (2018) and Hortigüela-Alcalá et al., (2015a), in which the highest scores were given by the students of later years. A possible explanation could be that students already knew these advantages, so they did not appreciate them as much as the previous year. Another possibility is that the different group compositions (part of the students were enrolled in the double degree), as well as their personal traits, affected this variable (Nauzer & Jaunky, 2021). It could also be that the workload due to the course organisation would have led to slightly lower scoring of the advantages. Further research on these aspects seems necessary, as well as to explore how this workload is managed (Schully & Kerr, 2014) and what exactly it refers to, since students' perception of their own workload may be different from the teachers' perception (Alshamy, 2017) and may be influenced by factors such as content, difficulty, type of assessment or even teacher-student relationships (Kember, 2006).

The results found with regard to the FSA system disadvantages are very disparate and they can be divided into three groups: items with very high, average and average-low scores. The items with high scores could actually be considered as features of the FSA system (attendance and active participation are mandatory, it requires participating

ongitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

in my own assessment, it requires continuity and it is more demanding), similar to what Gallardo et al. (2018), López-Pastor (2009) and Martínez-Mínguez et al. (2015) found in their research. The high score on 'the working method is not very well-known' could mean that it is not common to find FSA systems in these degree studies.

Among the disadvantages, five items yielded statistically significant differences between the two courses: (a) the working method is not very well-known, lack of habit; (b) it requires continuity; (c) it is more demanding; (d) there is a mismatch between work and credits; (e) it is unfair, compared to other assessment processes. These values were always lower in the second course. It seems like students deemed the potential disadvantages as less important as they progressed within the degree and gained experience with the FSA system. This aspect was also brought to light by Gallardo et al. (2018) in their study about PSTE-EC, where students gave lower scores to the disadvantages in the later course. The data seem to indicate that the more the students learned to work with the FSA system, the lower the importance they gave to the potential disadvantages and the lower the workload needed. This could mean that they were learning how to generate learning self-regulation processes (Panadero et al., 2014).

Although statistically significant differences were not found in academic performance between groups, the results revealed high academic performance, which agrees with previous studies addressing the influence of FSA on this variable (Hortigüela-Alcalá et al., 2019; Martínez-Mínguez et al., 2019). Panadero et al. (2014) found that student participation in the assessment process was a determining factor in academic performance, due to students' self-regulation and involvement in their own work. When comparing academic performance between the two courses, the 'pass' percentage and the marks obtained were higher in the second course, but the differences were not statistically significant. Therefore, we cannot state that the FSA system leads to higher academic performance as experience increases. However, knowing the FSA system used seemed to help generate learning self-regulation processes, as students are conscious of their learning process and how they need to regulate their tasks (Zapana, 2019).

Strictly speaking, the hypothesis proposed cannot be accepted because, although clear differences were observed (2.80 vs. 3.19), they were not statistically significant. This is in contrast with the study by Hortigüela-Alcalá et al., (2015b) where, after having previously experienced FSA systems, students showed more confidence, knew the process and were usually more involved, which helped to improve their academic performance, since this experience also provided them with more resources and skills to solve tasks.

Besides, the majority of students chose the continuous type of assessment in both courses, as happened in similar studies (Castejón-Oliva et al., 2011; Fraile-Aranda et al., 2013). When comparing the assessment type choice, students showed an evolution towards greater involvement in the second course. Thus, the percentage of students who chose the continuous and mixed types increased slightly. This may be because they already knew the FSA system and they were aware that they would learn much less and would, in general, achieve a very low academic performance with the 'final' assessment type.

If we combine assessment types and academic performance, we can see that the latter was clearly higher with continuous assessment in both courses: (a) Distinction and Outstanding were only obtained through this assessment type in both courses; and (b) the students who chose the final assessment type showed a much lower academic performance, with a higher percentage of 'fail' than with the continuous or mixed types. Similar findings were reported by Fraile-Aranda et al. (2013) and Romero-Martín et al. (2014), who confirmed that those students who chose the final assessment type achieved lower performance than those who were engaged in formative types, such as continuous or mixed.

A possible explanation for this low academic performance is that it is complicated for those students who do not attend the lessons or complete the proposed learning tasks to acquire the competencies or knowledge that will later be assessed in the course. Similar results were reported by Castejón et al (2011), López (2009) and López et al (2013). These studies indicated that the mixed type could lead to very variable academic performance results, depending on the students' level of involvement and learning.

With regard to teacher workload, the most timeconsuming task was correcting documents. The implementation of the FSA system requires providing regular feedback to students (Gutiérrez-García et al., 2011; Romero-Martín et al., 2014). The comparison between courses was remarkable: despite the same FSA system having been used in both of them and both having the same number of ECTS credits, the workload was considerably higher in the second one (164 vs. 109 h). We believe this could be due to several factors: (a) a higher number of students; (b) the number of teachers in each course; and (c) the shorter experience of two of the teachers. The number of students did not completely explain the difference, since the mean student workload increased from 2.6 to 3.15 hours. It could be because the first course was taught by an experienced teacher (12 years teaching this course), while the second one was taught by three teachers, two of them with little teaching experience. In this regard, López et al. (2013) discovered that, in courses where an FSA system was used, the teacher workload depended not only on the number of students, but also

on other factors like, for example, the learning activities conducted during the course.

Moreover, the student workload was very similar to the number of non-lesson hours corresponding to the course credits (90 hours). This adaptation of the FSA systems to the actual workload with ECTS credits has previously been verified (López et al., 2013; Martínez-Mínguez et al., 2015; Romero-Martín et al., 2015), refuting that the implementation of FSA systems leads to student overload. Nonetheless, it is true that it entails a greater workload than the traditional method, where students usually only study for a few weeks prior to the final exam (Hortigüela-Alcalá et al., 2015b)

With regard to student workload, significant differences were detected in the total number of hours spent on each course (96 vs. 85), with α <.001. The student workload was higher in the first course. This seems to indicate that a better knowledge of FSA helped to spend a shorter time reviewing and remaking projects based on the feedback received. Another possible explanation is that they showed better self-regulation since they already knew the method and the teacher's demands; this was not due to general, but contextual self-regulation. This has been described in the literature as non-transfer from one context to another (Greene, 2020; Zakszeski et al., 2020).

There seemed to be a clear relationship between academic performance and student workload in both courses, except for those students who received Distinction in the first course, who showed a lower workload than those with Outstanding (115 vs. 150 h). There was no clear reason to explain this. It could be due to the competencies, intelligence, skills or knowledge these students already had or acquired, or their self-regulation ability. The results showed a significant positive relationship between student workload and academic performance. The relationship seemed to be stronger in the second course, with a lower workload needed to achieve the same results. This seems to indicate that students became more efficient once they knew and had experience with the assessment system. However, other factors must be borne in mind, such as students' self-regulation ability (Karaman, 2021). Díaz et al. (2016) observed that some qualitative aspects, such as time-planning skills, also had an influence on academic performance. Garbanzo (2007) mentioned other factors that affect academic performance in higher education: cognitive competence, motivation, social or socio-economic differences, relationships among students, course complexity, etc.

Conclusions

According to the results, the different variables analysed showed different evolutions. Some variables improved when the same FSA system was implemented in two supplementary courses taught in consecutive PSTE years,

but others did not, or did not clearly. Students deemed FSA advantages as positive, although some of them received slightly lower scores in the second course. The disadvantages received disparate scores (high, average or low) but, in contrast, a positive evolution was observed in the second course (lower scores for some disadvantages).

As regards academic performance, the success rate (percentage of 'pass') was high in both courses and better in the second one, but with no significant differences. Consequently, the data seem to indicate that the FSA system had a positive effect on academic performance.

The majority of students chose the continuous type of assessment. Moreover, this was the assessment type that showed the highest academic performance, followed by the mixed assessment. The final assessment type presented very low academic performance.

The results uncovered that the teacher workload was higher in the second course, which seemed to be due to the larger number of students and the teachers' more limited experience. Nevertheless, the student workload was well in accordance with the relationship between credits and hours of independent work established by the ECTS system. In fact, it was lower in the second course, as a logical effect of the students' learning self-regulation process. A clear significant relationship was observed between student workload and academic performance, with a strong difference between courses, since students seemed to need a lower workload in the second course to achieve similar performance.

No research was found addressing the impact of FSA on the same cohort of students during two different academic years. Therefore, this study makes a relevant contribution regarding the outcomes and benefits that FSA systems seem to generate within PSTE, in particular, on the disadvantages, academic performance and workload. It may be of interest to university teachers who would like to implement FSA systems in their lessons in the long term or to those who would like to get familiar with the use of FSA.

There were three major limitations: 1) in the second course, there were students from two different time contexts and this could be an uncontrolled extraneous variable; 2) the sample was small and it was collected from one single centre, so it would be advisable to replicate the study with larger samples and to compare them to other university contexts and degrees; and 3) the combination with a qualitative approach would help with result interpretation.

Future research lines could compare the outcomes of FSA systems to other cohorts of students, either from the same or different contexts. They can also be compared to other more traditional assessment systems or to other courses, and mixed designs could be applied to allow for better interpretation of certain aspects that could not be explored in this research.

Funding

Grant RTI2018-093292-B-I00 funded by MCIN/ AEI/10.13039/501100011033 and, by "ERDF A way of making Europe".

References

- Alshamy, A. (2017). Credit hour system and student workload at alexandria university: a possible paradigm shift. *Tuning Journal For Higher Education*, *4*(2), 277-309. https://doi.org/10.18543/tjhe-4(2)-2017pp277-309.
- Atienza, R., Valencia-Peris, A., Martos-García, D., López-Pastor, V. M. & Devís-Devís, J. (2016). La percepción del alumnado universitario de Educación Física sobre la Evaluación Formativa: ventajas, dificultades y satisfacción. *Movimento. Revista de Educação Física Da UFRGS*, 22(4), 1033-1048.
- Aubert, A., Flecha, A., García, C., Flecha, R., & Racionero, S. (2008). *Aprendizaje dialógico en la sociedad de la información*. Hipatia.
- Barrientos, E. J., López-Pastor, V. M. & Pérez-Brunicardi, D. (2019). ¿Por qué hago evaluación formativa y compartida y/o evaluación para el aprendizaje en educación física? La influencia de la formación inicial y permanente del profesorado. *Retos*, (36), 37-43. https://doi.org/10.47197/retos.v36i66478.
- Basera, C. H. (2019). Learners' perceptions of assessment strategies in higher education. *Journal of Education and e-Learning Research*, 6(2), 76-81. https://doi.org/10.20448/journal.509.2019.62.76.81.
- Castejón-Oliva, F., López-Pastor, V., 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 del Deporte, 11(42), 328-346. https://bit.ly/2Nb2vi8.
- Castejón-Oliva, F.J., Santos-Pastor, M. L., & Palacios, A. (2015). Cuestionario sobre metodología y evaluación en formación inicial en educación física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte, 15*(58), 245-267. https://doi.org/10.15366/rimcafd2015.58.004.
- Córdoba, T., et al. (2016). Educación física cooperativa, formación permanente y desarrollo profesional. De la escritura colectiva a un relato de vida compartido. *Retos*, *29*, 264-269. http://www.redalyc.org/html/3457/345743464049.
- Díaz, A., Lobos, K. & Bruna, D. (2016). *Fomento de la autorregulación del aprendizaje en el aula universitaria*. Universidad de Concepción.
- Dochy, F., Segers, M., & Dierick, S. (2002). Nuevas vías de aprendizaje y enseñanza y sus consecuencias: una era de evaluación. *Red Estatal de Docencia Universitaria*, *2*(2), 13–30. https://bit.ly/2ALg19m.

- Fraile-Aranda, A., López-Pastor, V. M., Castejón-Oliva, F. J., & Romero-Martín, R. (2013). La evaluación formativa en docencia universitaria y el rendimiento académico del alumnado. *Revista Aula Abierta, 41*(2), 23-34. https://bit.ly/3yTdTbV.
- Gallardo-Fuentes, F., & Carter-Thuillier, B. (2016). La evaluación formativa y compartida durante el practicum en la formación inicial del profesorado: análisis de un caso en Chile. Retos. Nuevas tendencias en Educación Física Deportes y Recreación, 29, 258-263. https://bit.ly/2YWZdVd.
- Gallardo, F., López-Pastor, V. M., & Carter, B. (2018). Efectos de la aplicación de un sistema de evaluación formativa en la autopercepción de competencias adquiridas en formación inicial del profesorado. *Revista Estudios Pedagógicos, 44*(2), 55-77. https://doi.org/10.4067/50718-07052018000200055.
- Garbanzo, G. M. (2007). Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública. *Revista Educación*, *31*(1), 43-63. https://bit.ly/38HgoOW.
- Greene, J. A. (2020). Building upon synergies among self-regulated learning and formative assessment research and practice. Assessment in Education: Principles, Policy & Practice, 27(4), 463-476. https://doi.org/10.1080/0969594X.2020.1802225.
- Gutiérrez-García, C., Pérez-Pueyo, A., Pérez-Gutiérrez, M., & Palacios-Picos, A. (2011). Percepciones de profesores y alumnos sobre la enseñanza, evaluación y desarrollo de competencias en estudios universitarios de formación de profesorado. *Cultura y Educación*, 23(4), 499-514. https://doi.org/10.1174/113564011798392451.
- Hortigüela-Alcalá, D., Palacios, A. & López, V. M. (2019). The impact of formative and shared or co-assessment on the acquisition of transversal competences in higher education. *Assessment & Evaluation in Higher Education,* 44(6), 933-945. https://doi.org/10.1080/02602938.2018.1530341.
- Hortigüela-Alcalá, D., Pérez-Pueyo, A. & Abella. V. (2015a). ¿De qué manera se implica el alumnado en el aprendizaje? análisis de su percepción en procesos de evaluación formativa. *Revista de Investigación en Educación, 13*(1), 88-104. https://bit.ly/3dnTE7n.
- Hortigüela-Alcalá, D., Pérez-Pueyo, A. & López-Pastor. V. M. (2015b). Implicación y regulación del trabajo del alumnado en los sistemas de evaluación formativa en educación superior. *RELIEVE*, *21*(1), 1-15. https://doi.org/10.7203/relieve.21.1.5171.
- Hortigüela-Alcalá, D., Pérez-Pueyo, A., & Fernández-Río, J. (2016). Influencia de las experiencias vivenciadas por el alumnado en el desempeño de futuras competencias docentes. *Contextos Educativos*, 19, 25-41. https://doi.org/10.18172/con.2742.

Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

- Karamar, P. (2021). The impact of self-assessment on academic performance: a meta-analysis study. *International Journal of Research in Education and Science* (*IJRES*), 7(4), 1151-1166. https://doi.org/10.46238/ijres.2344.
- Kember, D. (2004). Interpreting student workload and the factors which shape student's perceptions on their workload. *Studies in Higher Education*, *29*(2), 165-184 https://doi.org/10.1080/0307507042000190778.
- Knight, P. (2005). El profesorado de Educación Superior. Formación para la excelencia. Narcea.
- López-Pastor, V. M. (Coord.) (2009). La Evaluación Formativa y Compartida en Educación Superior: propuestas, técnicas, instrumentos y experiencias. Narcea.
- López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A. & Macdonald, D. (2013). Alternative assessment in physical education: a review of international literature. *Sport, Education and Society,* 1-20. https://doi.org/10.1080/13573322.2012.713860.
- López-Pastor, V. M., 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(2), 163-180. https://doi.org/10.1080/030 9877X.2011.644780.
- López, V. M., Molina, M., Pascual, C. y Manrique, J. C. (2020). La importancia de utilizar la evaluación formativa y compartida en la formación inicial del profesorado de educación física: los proyectos de aprendizaje tutorado como ejemplo de buena práctica. *Retos, 37*, 620-627. https://bit.ly/2Qhf6lA.
- López, V. M., Monjas, R., Manrique, J. C., Barba, J. J. & González, M. (2008). Implicaciones de la evaluación en los enfoques de educación física cooperativa. El papel de la evaluación formativa y compartida en la necesaria búsqueda de coherencia. *Cultura y Educación, 20*(4), 457-477. https://bit.ly/2YWCN6p.
- López-Pastor, V. M. & Sicilia-Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education*, 42(1), 77-97. https://doi.org/10.1080/02602938.2015.1083535.
- Martínez-Mínguez, L. & Flores, G. (2014). Profesorado y egresados ante los sistemas de evaluación del alumnado en la formación inicial del maestro de educación infantil. *RIDU*, 8(1). https://bit.ly/2zSXx6s.
- Martínez, L., Martín, M. & Capllonch, M. (2009). Una experiencia de desarrollo profesional del docente universitario de educación física a través de una

- práctica crítica, reflexiva y colaborativa. *Cultura y Educación*, *21*(1), 95-206. https://bit.ly/3fQnlya.
- Martínez-Mínguez, L., Moya, L, Nieva, C., & Cañabate, D. (2019).

 Percepciones de estudiantes y docentes: evaluación formativa en proyectos de aprendizaje tutorados.

 Revista Iberoamericana de Evaluación Educativa, 12(1), 59-84. https://doi.org/10.15366/riee2019.12.1.004.
- Martínez, L., Vallés, C. & Romero, R. (2015). Estudiantes universitarios: ventajas e inconvenientes de la evaluación formativa. *@tic revista d'innovació educativa, 14*, 59-70. https://doi.org/10.7203/attic.14.4217.
- Mérida, R. (2006). La convergencia europea y la formación universitaria en competencias para la docencia en Educación Infantil. *Revista de Educación, 341,* 663-686. https://bit.ly/3fQHPs3.
- Montero, I. & León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862. https://bit.ly/2YXjbPv.
- Nauzeer, S. & Jaunky, V. C. (2021). A meta-analysis of the combined effects of motivation, learning and personality traits on academic performance. Pedagogical Research, 6(3), em0097. https://doi.org/10.29333/pr/10963.
- Panadero, E., Alonso-Tapia, J. A. & Huertas, J. A. (2014). Rubrics vs. self-assessment scripts: effects on first year university students' self-regulation and performance. *Infancia y Aprendizaje: Journal for the Study of Education and Development, 37*(1), 149-183. https://doi.org/10.1080/02103702.2014.881655.
- Panadero, E. & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice, 21*(2), 133-148. https://doi.org/10.1080.0969594X.2013.877872.
- Pascual-Arias, C., García-Herranz, S. & López-Pastor, V. M. (2019). What do preschool students want? The role of formative and shared assessment in their right to decide. *Cultura y Educación, 31*(4). https://doi.org/10.1080/11356405.2019.1656486.
- Pascual-Arias, C. & Molina, M. (2020). Evaluar para aprender en el prácticum: una propuesta de evaluación formativa y compartida durante la formación inicial del profesorado. *Publicaciones*, *50*(1), 183-206. https://doi.org/10.30827/publicaciones.v50i.15959.
- Romero-Martín, R., Castejón-Oliva, F. J. & López-Pastor, V. M. (2015). Divergencias del alumnado y del profesorado universitario sobre las dificultades para aplicar evaluación formativa. *RELIEVE*, *21*(1). https://doi.org/10.7203/relieve.21.1.5169.
- Romero-Martín, R., Fraile-Aranda, A., López-Pastor, V. M., & Castejón-Oliva, F. J. (2014). Relación entre sistemas de

Longitudinal Study on the Effects of Formative and Shared Assessment on Pre-Service Teacher Education

- evaluación formativa, rendimiento académico y carga de trabajo del profesor y del alumno en la docencia universitaria. *Revista Infancia y Aprendizaje, 37*(1), 16-32. https://doi.org/10.1080/02103702.2014.918818.
- Santos, J. M., Ortiz, E. & Marín, S. (2018). Índices de variación de la nota debidos a la evaluación continua. Contrastación empírica en la enseñanza universitaria. *Cultura y Educación, 30*(3), 491-527. https://doi.org/10.1080/11356405.2018.1488422.
- Schully, G. & Kerr, R. (2014). Student workload and assessment: strategies to manage expectations and inform curriculum development. *Accounting Education*, 23(5), 443-446. https://doi.org/10.1080/09639284.2014.947094.
- Silva Rodríguez, I., & López-Pastor, V.M.; (2015). ¿Cómo vive el alumnado la evaluación en formación inicial del profesorado? *@-tic revista d´innovación educativa*, *14*. https://doi.org/10.7203/attic.14.4171.
- Southcot, J. & Crawford, R. (2018). Building critically reflective practice in higher education students: employing auto-ethnography and educational connoisseurship in assessment. *Australian Journal of Teacher Education*, 43(5), 95-109. https://doi.org/10.14421/ajte.2018v43n5.7.
- Zakszeski, B., Hojnoski, R. L., Dever, B. L., DuPaul, G. J. & McClelland, M. M. (2020). Early elementary trajectories of classroom behavior self-regulation: prediction by student characteristic and malleable contextual factors. *School Psychology Review, 49*(2), 161-177. https://doi.org/10.1080/2372966X.2020.1717373.
- Zapana, L. J. (2019). Influencia de la evaluación formativa y compartida en el proceso de autorregulación de los estudiantes de la carrera profesional de educación primaria del instituto de educación superior pedagógico público Arequipa 2018. [Trabajo Fin de Grado para optar el Grado Académico de Maestra en ciencias: Educación con mención en Gestión y Administración Educativa, Universidad Nacional de san Agustín de Arequipa, Perú].