

Effects of the Implementation of Tutored Learning Projects and Formative and Shared Assessment Systems in Pre-Service Teacher Education

Miriam Molina Soria
University of Valladolid

David Hortigüela-Alcalá
University of Burgos

Víctor M. López-Pastor
University of Valladolid

Cristina Pascual-Arias
University of Valladolid

Carla Fernández-Garcimartín
University of Valladolid

The main purpose of Formative and Shared Assessment (F&SA) is to improve student learning, to improve and redirect teaching practice and to improve the teaching-learning process. This research has two objectives: (1) to analyse the results of the applicability and difficulty of the Good Practice of Tutored Learning Projects (TLP) in a Pre-service Teacher Education (PTE) subject; and (2) to examine the advantages and disadvantages of the F&SA system used in the same subject. For this purpose, a longitudinal study of a PTE subject is carried out over three academic years. The instruments used to collect information are: a structured report on Good Practice in university teaching and an anonymous questionnaire for students to evaluate the experience of Good Practice, which has been carried out and the subject's evaluation system. The results show the consistency of the course design over the years. Students are satisfied with the TLP experience because they consider it to be useful, effective, innovative, replicable and sustainable, as well as helping in the acquisition of competences.

Keywords: Tutored Learning Projects, Formative Assessment, Shared Assessment, Pre-service Teacher Education

INTRODUCTION

Despite the great progress in society in terms of the term assessment, there is still much to learn and change in education. The confusion between the term assessment and grading is constant (Palacios and López-Pastor, 2013). Assessment should be associated with learning (Chiappe et al., 2016), and grading is

a numerical mark at the end of a process (Hortigüela-Alcalá et al., 2019). Therefore, an assessment process should not imply a grade. In this sense, Hortigüela-Alcalá et al. (2019) state that: “*Everything that can be graded should be assessable but not everything that can be assessed has to be gradable*” (p.14).

In this sense, within Pre-Service Teacher Education (PTE), it is necessary to apply assessment models focused on generating student learning, away from constant grading. In this sense, Formative and Shared Assessment (F&SA) systems play a fundamental role in PTE.

Formative and Shared Assessment in Pre-Service Teacher Education

Firstly, it's necessary to understand the concept of Formative Assessment and Shared Assessment. Formative Assessment is an assessment process whose main objective is to improve student learning, teaching practice and teaching-learning processes (López-Pastor and Pérez-Pueyo 2017). Bennett (2011) states that formative assessment cannot be conceived as a test or as a process, but as a reflective integration of the process and the methodology or instrumentation designed. Likewise, Hidalgo (2021) understands Formative Assessment to be:

A strategy for the acquisition of new knowledge in a more proactive way and for the development of skills and attitudes on the part of the students, with the university teacher being the key player, as they are the active and conscious protagonists in the management of evaluation (p. 191).

Shared Assessment is a dialogical process between teachers and students on the assessment of learning and teaching-learning processes (López-Pastor, 2009; López-Pastor and Pérez-Pueyo, 2017). In contrast, Salinas (2002) defines Shared Assessment as the sharing and coordination of teachers on the assessment to be used in the classroom. However, Ibarra et al. (2013) state that Shared Assessment is the sharing of assessment work between teachers and students, which enables strategies such as self-assessment and co-assessment to be put into practice. These authors differentiate between the term co-assessment and peer-assessment, differentiating between who assesses (students, students and teachers) and the degree of responsibility and participation in the assessment process.

Parra and García-Martínez (2021) state that active assessments in which the student plays a leading role (as in Shared Assessment), have increasingly more presence than traditional assessments due to their high formative value.

The combination of Formative Assessment and Shared Assessment (F&SA) seeks to generate student learning, improve and reorient teaching practice and redirect the teaching-learning process, all by involving students in the assessment process through different techniques (López-Pastor and Pérez-Pueyo, 2017). Some of the participation techniques that make it up are self-assessment (carrying out a self-assessment); peer-assessment (which can be intra- or inter-group); self-assessment (assigning a numerical mark to oneself, which should be a consequence of self-assessment); and dialogue grading (a process of dialogue between students and teachers to adjust a final numerical mark based on a previous self-assessment). In this sense, Gutiérrez et al. (2018) define F&SA as a feedback process between all educational agents whose main objective is to generate student learning and improve teaching practice, without associating this process with grading.

Several studies (Cañadas, 2021; Dorit and Nirit, 2020; Estevan et al., 2018; Gallardo and Carter, 2016; Gallardo et al., 2020; Hamodi and López-Pastor, 2012; Hortigüela-Alcalá, et al., 2021; Pascual-Arias and Molina, 2019; Romero et al., 2014) claim that it is convenient to experiment with F&SA systems in PTE because, in addition to being a specific knowledge competence, students learn more and better and their interest and motivation increase. There is work that defends F&SA systems in PTE for several reasons:

- (1) Because it improves the acquisition of professional competences, thanks to the involvement of students in their assessment process (Castejón et al., 2018; Gallardo and Carter, 2016; Romero et al., 2016).
- (2) Because it improves student learning (Hortigüela-Alcalá et al., 2019; Martínez-Mínguez et al., 2019).

(3) Because it improves students' academic performance (Fraile et al., 2013; Mastagli et al., 2020; Panadero and Jonsson, 2013).

(4) Because students and teachers have a high degree of satisfaction (Atienza et al., 2016).

According to López-Pastor et al. (2021) the use of F&SA systems in PTE has a double meaning: (1) it is an assessment system that affects students' learning processes; and (2) assessment is a professional competence to be acquired during PTE. Furthermore, these authors list six reasons why it is important to develop F&SA systems in PTE:

(1) It serves to learn more and better.

(2) It helps to better develop many teaching competences.

(3) It encourages students to focus more on their learning process and to take responsibility for it.

(4) It tends to lead to more educational success and better academic performance.

(5) It is a basic competence of every teacher, which is usually acquired more through practical experimentation than through theoretical study.

(6) It facilitates a better transfer between what is learned in PTE and educational practice in schools (López-Pastor et al., 2021, p.38).

Gutiérrez et al. (2018) assert that F&SA is a feedback process between the educational agents that are part of the process, which focuses on improving student learning and teaching practice. Therefore, feedback is an important aspect of F&SA systems. Feedback is the information provided to students to correct and redirect their mistakes (Barrientos et al., 2019; Canabal and Margalef, 2017; Martínez, 2021; Nicol et al., 2014; Panadero et al., 2017; Saiz-Linares and Susinos-Rada, 2018). This information should not only be provided by the teacher, but these processes can be carried out with the participation of students as a group, for example, through self-assessment or peer-assessment (Hortigüela-Alcalá et al., 2019). In this sense, Bono and Núñez-Peña (2018) state that when students provide feedback to each other (they monitor, evaluate and regulate their learning), this has an impact on their learning. Furthermore, if they receive external feedback from the teaching staff, they argue that this information should be converted into internal feedback among students to generate this impact on learning, because when they give each other feedback among peers, at the same time they think, reflect and generate more feedback on their own work, which makes them more independent.

Parra and García-Martínez (2021) conducted a study with two groups to analyse the impact of teacher feedback on student evaluations. One group received feedback after the teacher evaluation and the other did not. The results show that both groups improved after the assessment, but there was a significantly greater improvement in the group that received feedback. Winstone and Boud (2021) point out a number of problems arising from the confusion between the terms assessment and feedback. They identify the main problems as being:

(1) Students focus on grades rather than developmental information.

(2) Teachers focus feedback on the justification for the mark rather than providing developmental information.

(3) Feedback information may come too late to be useful.

(4) Course design does not focus on feedback.

(5) The ways of documenting feedback may impair its quality.

(6) Anonymous grading depersonalises and inhibits the quality of feedback.

The authors also compile a collection of strategies to preserve the learning function of feedback:

(1) Adaptive release of grades.

(2) The elicitation of feedback from learners.

(3) Designing for the application of feedback information.

(4) Reallocating time spent providing detailed feedback to a different part of the learning cycle.

(5) Separate anonymous grading from personalised feedback.

(6) Encourage continuous feedback curation.

For all the above reasons, feedback should help learners to progress in their learning process, and can be given in written or oral form, before, during or after an activity.

However, according to various studies (López-Pastor and Pérez-Pueyo, 2017; Molina and López, 2017; Nieva et al., 2021; Martínez-Mínguez et al., 2015; Romero et al., 2015; Vallés et al., 2011), Table 1 shows the advantages and disadvantages of F&SA in PTE.

TABLE 1
ADVANTAGES AND DISADVANTAGES OF F&SA

Advantages	Disadvantages
<ul style="list-style-type: none"> – Active and meaningful learning takes place. – It helps to acquire personal and professional competences. – It encourages the monitoring of students on a more individual basis. – There is a relationship between theory and practice. – The student is a participant in his or her own evaluation process. – It improves academic performance. – Constant and quality feedback helps to improve student learning. 	<ul style="list-style-type: none"> – The workload is greater for both teachers and students. The degree of existence is greater than in other types of assessment. – It requires daily continuity and active class attendance. – The work dynamics may be unfamiliar due to lack of experience in F&SA.

Source: own elaboration (based on López-Pastor and Pérez-Pueyo, 2017; Molina and López, 2017; Nieva et al., 2021; Martínez-Mínguez et al., 2015; Romero et al., 2015; Vallés et al., 2011).

Tutored Learning Projects and Formative and Shared Assessment Systems

Tutored Learning Projects (TLP) are a work that combines a theoretical and a practical part on a topic (Barba-Martín et al., 2010; Gallardo et al., 2018; López-pastor et al., 2021). TLP have been used in PTE for several decades because they are a very effective methodology with good results in terms of learning and student satisfaction (Barba et al., 2010; Gallardo et al., 2020; López-Pastor et al., 2020, 2021; Martínez-Mínguez, 2016; Martínez-Mínguez et al., 2019). TLP are combined with F&SA systems because it seems to be the most logical assessment, as constant feedback during the development of the work helps to improve student learning and to obtain a good quality of the documents.

López-Pastor et al. (2020) present an experience combining TLP and F&SA in PTE. The results show that students do not consider the experience to be of high difficulty and that it is useful, innovative, effective and replicable. Along the same lines, López-Pastor et al. (2021) present a proposal for TLP and F&SA in which they detail the assessment instruments used during the process (graded scales, group self-assessment form and form for reflective assessment on the implementation of TLP) and explain how the feedback is carried out: in all the necessary deliveries and tutorials the teacher provides oral and written feedback on the documents, and the students have the possibility of improving until they reach an acceptable quality.

In other studies, students' perception of the combination of TLP and F&SA is that they are quite satisfied with the experience and the assessment system and, in addition, students perform well academically (Barba and López, 2017; Martínez-Mínguez et al., 2019). In this sense, Manrique et al. (2010) claim that this combination of TLP and F&SA also helps students to acquire professional competences.

Currently, there are few longitudinal studies that analyse the results of the applicability of TLP in the FIP in the same subject and with the same teaching staff; that is, there is little published data on this type of study in terms of the effects of competence acquisition, the evaluation and feedback system provided during the process or student satisfaction with the experience and the evaluation system used. Therefore, the aim of this research is to check the consistency of the application of the F&SA over three academic years. The objectives pursued are (1) to analyse the results of the applicability and difficulty of the TLP Good Practice, as well as competence acquisition, assessment, feedback and student satisfaction, combined

with the F&SA systems in a PTE subject over three academic years; and (2) to examine the advantages and disadvantages of the F&SA system used in the same subject.

MATERIAL & METHODS

Research Method and Sample

This is a longitudinal study with a mixed methodology (qualitative and quantitative) in which comparisons are made over three academic years of the same subject with different students (Montero and León, 2005). The same F&SA system is applied in all the courses with the same teachers and, in addition, the same learning activities are carried out. The aim is to check the reliability of applying TLP and F&SA over time.

This study is carried out in a PTE subject at a Spanish public university: *Children's Body Expression and Communication*. The sample used is 114 students in the 4th year of the Early Childhood Education Degree, distributed over three academic years (see Table 2).

TABLE 2
STUDY SAMPLE

Academic year	Nº of students enrolled	Gender	Average age
2019-2020	37	93.3% women 6.7% men	22,85 years
2020-2021	35	96.2% women 3.8% men	23,03 years
2021-2022	42	91.8% woman 8.2% men	22,82 years
TOTALS	114	100%	22,8 years

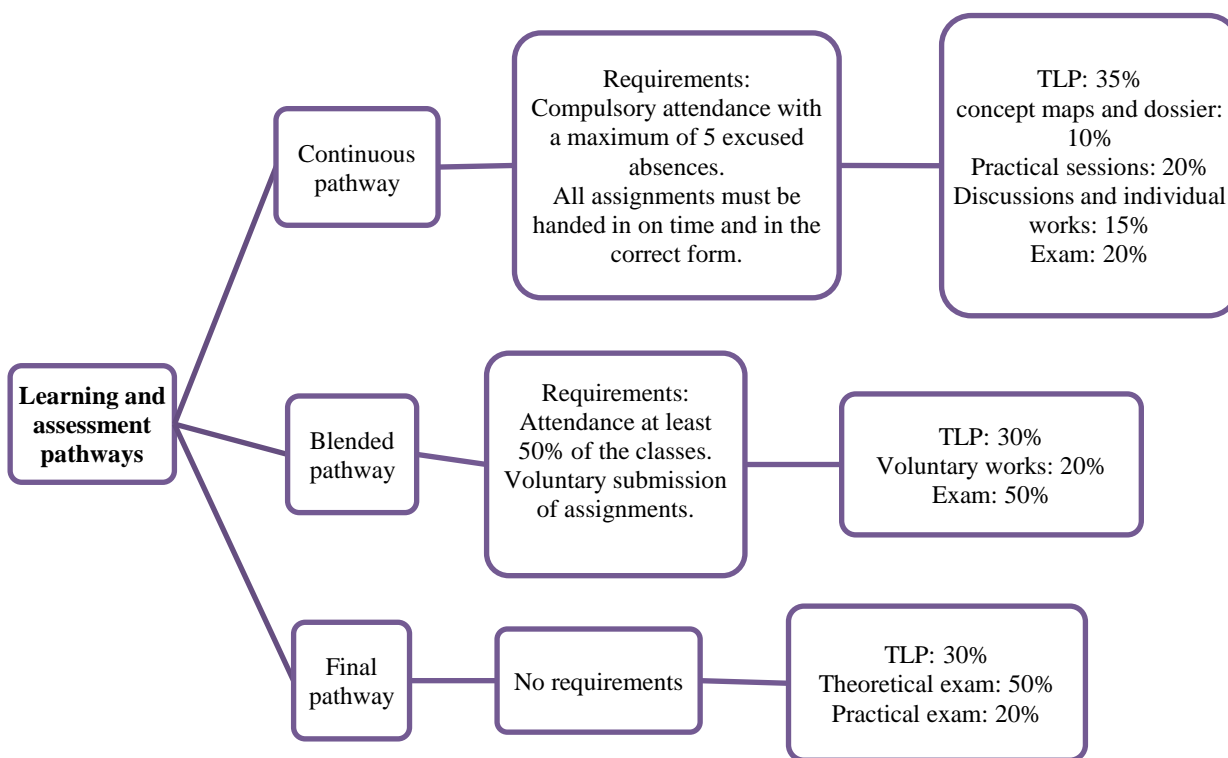
The teachers responsible for the subject are the same and the same F&SA system is used and the same TLP is carried out throughout all the courses.

On the first day of class, the assessment and grading criteria for the subject are agreed by consensus between students and teaching staff. Three learning and assessment pathways are offered in order to pass the subject. Figure 1 shows the requirements that must be met and the weight of each activity in the grade for each of the pathways.

The F&SA system developed in the course is based on constant feedback, both on the work and documents handed in by the students, as well as during the face-to-face classes. The assignments are corrected by the teachers, who provide the necessary feedback within one week. In the same way, students have one week to correct the documents and re-submit them if necessary.

Throughout the course, students are involved in their own assessment process; there is constant self-assessment and peer-assessment. In addition, at the end of the course, there is a process of self-assessment and dialogue grading to arrive at the final mark for the course. Throughout the four-month period, no numerical marks are given to the students; they work with rubrics and graded scales that the student has from the first day of class.

FIGURE 1
REQUIREMENTS FOR EACH OF THE PATHWAYS OFFERED IN THE SUBJECT AND THE WEIGHT OF THE LEARNING ACTIVITIES IN THE GRADE



The TLP is a learning activity carried out in the subject. As can be seen in Figure I, it is a compulsory activity, regardless of the learning and assessment pathway chosen by the student. It is a work that combines theory and practice. On the first day of class, groups of 3-4 people are formed and students are given a set of topics to work on. On the same day, a draw is made in which the groups of students choose the TLP topic. The steps to be followed in the elaboration of the TLP are:

- (1) Hold a tutorial with the tutor to provide the basic bibliography for the theoretical framework on the chosen topic. This document should occupy four pages. The teaching staff provides the necessary feedback in tutorials to correct the document until it is of good quality.
- (2) Design a practical session on the topic chosen for the early childhood education stage, but which they must put into practice in the faculty gymnasium with their classmates. In the same way, this session is corrected by the teaching staff through constant feedback until it is of good enough quality to be put into practice.
- (3) Putting the session into practice in the gymnasium and producing a final report detailing the changes that have taken place between what was planned and what happened, analysing the teaching competences and the data collected by the evaluation instruments they have designed for the implementation of the session. Once again, this document is corrected by the teaching staff providing feedback for good quality.

Students who experience the session that their classmates have designed must complete a practice sheet in which they analyse what happened in the session. In this document, a narration of the session, an analysis of the advantages and disadvantages of the practical proposal and of the TLP topic, a personal experience and an analysis of the teaching competences of the classmates who have presented the TLP are provided.

Data Collection Instruments

To collect the data for this study and to study the consistency of the application of TLP and F&SA systems over three academic years, the data collection instruments used in this study have been validated by the F&SA in Education Network:

- (1) A structured report of Good Practices in university teaching (Hortigüela-Alcalá et al., 2018). This report collects the subject data on the experience of Good Practice of TLP and the advantages and disadvantages of the assessment system that has been developed. This instrument is carried out at the end of the subject with the data from the anonymous questionnaire explained below.
- (2) An anonymous questionnaire for students to evaluate the Good Practice experience that has been carried out and the assessment system of the subject (validated by Castejón et al., 2015, with a validity of RMSEA= 0.078 and a reliability index of 0.84). The questionnaire has a Likert-type scale with five levels: 1 (not at all), 2 (a little), 3 (somewhat), 4 (quite a lot) and 5 (a lot). At the end of the questionnaire there is an open question for students to make comments if necessary. This questionnaire is completed on the last day of the course. This questionnaire aims to measure, according to the students' perception, (1) the applicability and difficulty of the TLP experience; (2) whether the TLP experience helps to acquire professional competences; (3) the F&SA system used and the feedback given during the process; and (4) student satisfaction.

Data Analysis

The data are analysed around three main blocks: (1) the results of the questionnaire on the students' perception of the implementation of the TLP; (2) the students' perception of the advantages of the F&SA system of the subject; and, (3) the disadvantages of the F&SA system.

As for the results of the questionnaire, the items have been divided into several categories:

- (1) Applicability and difficulty of the TLP experience.
 - Do you consider what you have learned from this experience to be useful?
 - It is an innovative experience, because it develops new or creative solutions.
 - It is an effective experience, because it demonstrates a positive and tangible impact of improvement.
 - It is a sustainable experience, because it is maintained over time and can produce long-lasting effects.
 - It is a replicable experience, when it is possible to use it as a model to develop it in other contexts.
 - What is the degree of difficulty of the experience?
- (2) Acquisition of competences
 - Do you think that this experience has helped you to acquire professional competences?
 - Does the evaluation that has been planned favour the acquisition of professional competences?
- (3) Assessment and feedback
 - How do you value the support received by the teacher?
 - How do you rate the help you received from your colleagues?
- (4) Student satisfaction.
 - Indicate overall satisfaction in relation to the experience.
 - Indicate the overall satisfaction in relation to the evaluation of the experience.

As for the data analysis, a descriptive quantitative analysis (arithmetic mean (\bar{X}), standard deviation (σ) and mode (M)) and an inferential analysis (one-factor ANOVA) are performed to check if there are significant differences ($p \leq 0.05$); in this way, we want to check if there is consistency in the application of the same F&SA system in the subject throughout the three academic years. In the one-factor ANOVA test, the dependent variables are the questionnaire items, while the independent variable is the academic year.

The open-ended question at the end of the questionnaire collects the observations of some students on their experience with TLP. All existing responses have been selected, as comments are scarce. The responses were categorised around the repetition of resources, the applicability of TLP in Early Childhood Education, the support provided by teachers and the lack of motivation of pupils.

RESULTS

This section analyses the results obtained. Firstly, the results of the TLP questionnaire are presented, then the advantages of the F&SA system used in the subject and, finally, its disadvantages. Table 3 shows the results of the descriptive analysis of the TLP questionnaire data over the three academic years.

TABLE 3
RESULTS OF THE DESCRIPTIVE ANALYSIS OF THE TLP QUESTIONNAIRE DATA OVER THE THREE ACADEMIC YEARS

Items	Course 2019-2020			Course 2020-2021			Course 2021-2022		
	\bar{X}	σ	M	\bar{X}	σ	M	\bar{X}	σ	M
Applicability and difficulty of the TLP experience									
Do you consider what you have learned from this experience to be useful?	4.48	.566	5	4.70	.535	5	4.63	.490	5
It is an innovative experience, because it develops new or creative solutions.	4.27	.719	4	4.10	.885	5	4.20	.632	4
It is an effective experience, because it demonstrates a positive and tangible impact of improvement.	4.64	.549	5	4.43	.568	4	4.34	.591	4
It is a sustainable experience, because it is maintained over time and can produce long-lasting effects.	4.30	.648	4	4.33	.802	5	4.18	.673	4
It is a replicable experience, because it can be used as a model to be developed in other contexts.	4.58	.561	5	4.60	.675	5	4.44	.561	4
What is the degree of difficulty of the experience?	3.15	.870	3	2.97	.669	3	3.40	.812	3
Skills acquisition									
Do you think that this experience has helped you to acquire professional competences?	4.73	.452	5	4.67	.479	5	4.54	.505	5
Does the assessment that has been planned favour the acquisition of professional competences?	4.33	.645	4	4.53	.730	5	4.20	.677	4
Assessment and feedback									
How do you value the help received by the teacher?	4.48	.755	5	4.30	.651	5	4.17	.707	4
How do you value the help received by your colleagues?	3.76	.830	4	4.17	.791	5	3.89	.963	4
Satisfaction									
Indicates the overall satisfaction in relation to the experience .	4.34	.653	4	4.43	.568	4	4.11	.631	4
Indicates overall satisfaction in relation to the evaluation of the experience.	4.03	.637	4	4.10	.548	4	3.74	.701	4

According to the results of Table 3 regarding the category of applicability and difficulty of the TLP experience, students consider TLP to be useful, innovative, effective, sustainable and replicable (means between 4.18 and 4.70). In addition, students do not perceive the experience as very difficult (averages between 2.97 and 3.40).

As for the second category concerning the acquisition of competences, the students agree quite strongly that the TLP help to acquire professional competences (averages between 4.54 and 4.73), and that the evaluation system used was quite conducive to the acquisition of these competences (averages between 4.20 and 4.53).

With regard to assessment and feedback, the results show that the help received by the teacher has a high average (between 4.17 and 4.48), but the help received by classmates has a lower average (between 3.76 and 4.17), although it is still a high average. Thus, students rate teaching help better than help from their peers.

Student satisfaction reflects that they are quite satisfied with the TLP experience. Moreover, in relation to the evaluation of the experience, the students are somewhat and fairly in agreement with the F&SA system used (averages between 3.74 and 4.10).

With regard to the open question at the end of the questionnaire, "Do you have any comments or observations about the experience," some students wanted to reflect on various issues:

Regarding the repetition of resources, some students consider that:

"We have seen repetitive things and some of them are difficult to carry out in our future teaching. Due to the lack of willingness of many centres, the lack of innovation and time raised." (A25- 2019/20)

In contrast, other students point out that, with adaptations, TLP resources can be carried out with children:

"Lots of resources to use (being able to be modified and adapted) tomorrow with the little ones." (A14- 2019/20)

Pupils are grateful for the help received by the teaching staff when the workload accumulates:

"I would like to comment that the teacher has been quite good in giving us time to do the PAT even though we were behind schedule." (A5- 2020/21)

Finally, one group points out the lack of motivation and willingness to learn throughout the implementation of the different TLP groups:

"To my TLP group it is a challenge that we try to learn and enjoy both ourselves and pass it on to the rest, however, there are many people who do not take it that way and it makes the TLP practices a bit heavy at times." (A30- 2019/20)

In the following Table 4, we present the results after performing the one-factor ANOVA test. In our case, the independent variable is the academic year, since we want to know if there are significant differences between the groups of the different academic years. The significance value is $p \leq 0.05$.

TABLE 4
RESULTS OF THE ONE-FACTOR ANOVA TEST OF THE ITEMS REFERRING TO TLP

Questionnaire items	Sum of Squares	df	Mean Square	F	ANOVA
1. Do you think that this experience has helped you to acquire professional competences?	.602	2	.301	1,306	.276
2. Does the evaluation that has been proposed favour the acquisition of professional competences?	1.804	2	.902	1.930	.151
3. Do you consider what you have learnt from this experience useful?	.766	2	.383	1.362	.261
4.1. It is an innovative experience, because it develops new or creative solutions.	.471	2	.235	.423	.656
4.2. It is an effective experience, because it demonstrates a positive and tangible impact of improvement.	1.519	2	.760	2.337	.102
4.3. It is a sustainable experience, because it is maintained over time and can produce long-lasting effects.	.453	2	.227	.439	.646
4.4. It is a replicable experience, when it is possible to use it as a model to develop it in other contexts.	.481	2	.240	.672	.516
5. How do you value the help received by the teacher?	1.680	2	.840	1.680	.192
6. How do you value the help received by your colleagues?	2.730	2	1.365	1..807	.170
7. State your overall satisfaction with the experience.	1.789	2	.895	2.328	.103
8. State your overall satisfaction in relation to the evaluation of the experience.	2,389	2	1,195	2,959	.057
9. What is the degree of difficulty of the experience?	3,085	2	1,542	2,458	.091

As can be seen in Table 4, there are no significant differences in the means of the academic years studied. This reflects the consistency of the TLP design and the F&SA systems of the subject throughout the academic years.

Table 5 presents the results of the students' perception of the advantages of the F&SA system implemented in the subject.

TABLE 5
STUDENTS' PERCEPTION OF THE ADVANTAGES OF THE SUBJECT'S F&SA SYSTEM

Items	Course 2019-2020		Course 2020-2021		Course 2021-2022				
	\bar{X}	σ	M	\bar{X}	σ	M	\bar{X}	σ	M
1. Offers alternatives to all students.	4.34	.653	4	4.40	.621	4	4.23	.808	5
2. There is a prior, negotiated and consensual contract of the evaluation system.	4.22	1.039	4	4.43	.898	5	4.50	.615	5
3. It is focused on the process, importance of daily work.	4.50	.672	5	4.90	.305	5	4.51	.562	5
4. The student is an active learner.	4.75	.440	5	4.87	.346	5	4.46	.657	5
5. Teamwork is approached in a collaborative way.	4.59	.550	5	4.70	.466	5	4.60	.553	5
6. The learner is more motivated, the learning process is more motivating.	4.03	.752	4	4.30	.596	4	4.06	.765	4
7. Grading is fairer.	3.94	.669	4	4.27	.691	4	3.86	.845	4
8. Improved academic tutoring (monitoring and helping the student).	4.50	.762	5	4.23	.568	4	4.26	.553	4
9. Allows functional learning.	4.50	.508	4	4.37	.556	4	4.40	.710	4
10. Generates meaningful learning.	4.63	.492	5	4.73	.450	5	4.40	.553	4
11. Much more is learned.	4.53	.621	5	4.70	.596	5	4.34	.684	5
12. Improves the quality of the work required.	4.44	.716	5	4.53	.681	5	4.09	.818	4
13. There is an interrelation between theory and practice.	4.72	.523	5	4.53	.629	5	4.49	.658	5
14. It assesses all possible aspects (in terms of knowledge, know-how, being and being).	4.50	.508	4	4.53	.507	5	4.29	.519	4
15. There is feedback on documents and activities.	4.69	.535	5	4.70	.535	5	4.43	.608	5
16. There is a possibility to correct mistakes in documents and activities.	4.84	.369	5	4.83	.461	5	4.51	.562	5
17. There is more individualised follow-up.	4.19	.780	4	4.17	.913	5	3.97	.747	4
18. Requires more accountability.	4.65	.608	5	4.63	.556	5	4.63	.598	5

As can be seen in Table 5, the results are very positive: the arithmetic means are in all courses higher than 3.86 for all items.

The items with the highest arithmetic mean across the three courses, and with a mode of 5 are: 3. *It is focused on the process, importance of daily work*; 4. *There is an interrelation between theory and practice*; 15. *There is feedback on documents and activities*; 16. *There is a possibility to correct mistakes in documents and activities*; and 18. *It requires more responsibility*.

Therefore, students consider these to be the main advantages of the F&SA system implemented in the subject (averages between 4.34 and 4.90).

The item with the lowest mean across the three courses is *The grading is fairer*, with means between 3.86 and 4.27. But it still has quite high averages. So, in general, students agree quite well with the items analysed.

Table 6 shows the results after the one-factor ANOVA test. Similarly, the independent variable is the academic year in order to find out whether there are significant differences between the different groups. The significance value is $p \leq 0.05$.

TABLE 6
RESULTS OF THE ONE-FACTOR ANOVA TEST OF THE ITEMS REFERRING TO THE ADVANTAGES OF THE F&SA SYSTEM USED IN THE SUBJECT

Questionnaire items	Sum of Squares	df	Mean Square	F	ANOVA
1. It offers alternatives to all students.	.503	2	.251	.507	.604
2. There is a prior, negotiated and consensual contract on the evaluation system.	1.404	2	.702	.942	.394
3. It is focused on the process, importance of daily work.	3.196	2	1.598	5.474	.006*
4. The student is an active learner.	2.940	2	1.470	5.722	.005*
5. There is a collaborative approach to team work.	.220	2	.110	.392	.677
6. The student is more motivated, the learning process is more motivating.	1.347	2	.673	1.328	.270
7. Grading is fairer.	2.962	2	1.481	2.676	.074
8. It improves academic tutoring (monitoring and help for the student).	1.349	2	.675	1.426	.245
9. It allows functional learning.	.303	2	.152	.521	.596
10. It allows meaningful learning.	1.903	2	.952	3.764	.027*
11. Much more is learned.	2.072	2	1.036	2.553	.083
12. It improves the quality of the work required.	3.689	2	1.844	3.329	.040*
13. There is an interrelation between theory and practice.	.992	2	.496	1.344	.266
14. It evaluates all possible aspects (knowing, knowing how to do, knowing how to be or being).	1.205	2	.602	2.301	.106
15. There is feedback on documents and activities.	1.573	2	.787	2.486	.089
16. There is the possibility of correcting mistakes in documents and activities.	2.356	2	1.178	5.242	.007*
17. There is a more individualised follow up.	.956	2	.478	.725	.487
18. It requires more responsibility.	.005	2	.002	.007	.993

As can be seen in Table 6, there are significant differences between the averages of the different academic years in five items: 3.- *It is focused on the process, importance of daily work*; 4.- *The student carries out active learning*; 10.- *It allows significant learning*; 12.- *It improves the quality of the work required*; and 16.- *There is the possibility of correcting errors in documents and activities*.

Table 7 shows the results of the students' perception of the disadvantages of the F&SA system of the subject.

TABLE 7
STUDENTS' PERCEPTION OF THE DRAWBACKS OF THE F&SA SYSTEM OF
THE SUBJECT

Items	Course 2019-2020			Course 2020-2021			Course 2021-2022		
	\bar{X}	σ	M	\bar{X}	σ	M	\bar{X}	σ	M
1. It requires compulsory and active attendance.	4.56	.759	5	4.63	.556	5	4.09	1.011	4
2. It has an unfamiliar working dynamic, lack of habit.	2.68	1.275	2	3.23	1.633	3	2.60	1.265	3
3. It requires continuity.	4.72	.457	5	4.77	.504	5	4.69	.631	5
4. It is necessary to understand it beforehand.	4.16	.847	5	4.13	.681	4	4.06	.838	4
5. Requires more effort.	4.44	.914	5	4.30	1.149	5	4.40	.812	5
6. It is difficult to work in a group.	2.72	1.301	2	2.43	1.165	2	2.79	.946	2
7. A lot of work can pile up at the end.	3.77	1.383	5	3.93	1.143	5	3.54	1.197	5
8. There is a disproportion between work/credits.	3.88	1.129	5	2.87	1.676	1	3.57	1.170	3
9. The grading process is more complex and sometimes unclear.	2.16	1.036	1	2.77	1.524	2	2.66	1.027	2
10. It generates insecurity and uncertainty, doubts about what is to be done.	2.06	.914	1	2.53	1.008	3	2.46	1.172	2
11. It is unfair compared to other evaluation processes.	2.03	.933	1	2.07	1.461	1	2.37	1.239	1
12. The corrections have been unclear.	1.56	.948	1	2.40	1.276	1	1.94	1.187	1
13. The assessment of the work is subjective.	2.00	1.136	1	2.27	1.143	1	2.44	1.501	1
14. It requires me to participate in my own evaluation (self-evaluation).	4.50	.672	5	4.50	.777	5	4.51	.742	5

As Table 7 shows, the items with the highest arithmetic mean (means between 4.06 and 4.72) and which, therefore, are the main disadvantages according to the students' perception are: 1; *Requires compulsory and active attendance*; 3. *It requires continuity*; 4. *It is necessary to understand it beforehand*; 5. *It requires more effort*; and 14. *It requires me to participate in my own evaluation (self-evaluation)*. Items 3, 5 and 14 always have a mode of 5, so they always have a very high rating.

As for the items with the lowest arithmetic mean, and which, therefore, students do not consider disadvantages of the F&SA system because they have little or no agreement with them, they are: 2. *it has unfamiliar work dynamics, lack of habit*; 6. *there is difficulty in working in groups*; 9. *The marking process is more complex and sometimes unclear*; 10. *It generates insecurity and uncertainty, doubts about what to do*; 11. *It is unfair compared to other assessment processes*; and 12. *Corrections have been unclear*; and 13. *The assessment of the work is subjective*.

This means that students do not consider these items to be drawbacks of the F&SA system used in the subject. In addition, the item corrections were unclear is the item with the lowest arithmetic means (between 1.56 and 2.40).

Table 8 presents the results of the one-factor ANOVA test on the items of the questionnaire on the drawbacks of the F&SA system carried out in the subject. Likewise, the independent variable is the academic year in order to find out if there are significant differences between the different groups. The significance value is $p \leq 0.05$.

TABLE 8
RESULTS OF THE ONE-FACTOR ANOVA TEST OF THE ITEMS REFERRING TO THE DISADVANTAGES OF THE F&SA SYSTEM USED IN THE SUBJECT

Questionnaire items	Sum of Squares	df	Mean Square	F	ANOVA
1. It requires compulsory and active attendance.	5.921	2	2.960	4.518	.013*
2. It has an unfamiliar work dynamic, lack of habit.	7.449	2	3.724	1.918	.153
3. It requires continuity.	.106	2	.053	.182	.834
4. 4.- It is necessary to understand it beforehand.	.181	2	.091	.143	.867
5. It requires a greater effort.	.312	2	.156	.169	.845
6. It is difficult to work in a group.	2.262	2	1.131	.866	.424
7. A lot of work can be accumulated at the end.	2.518	2	1.259	.813	.447
8. There is a disproportion between work and credits.	16.503	2	8.252	4.630	.012*
9. The marking process is more complex and sometimes not very clear.	6.460	2	3.230	2.218	.115
10. It generates insecurity and uncertainty, doubts about what has to be done.	4.055	2	2.028	1.868	.160
11. It is unfair in comparison with other assessment processes.	2.354	2	1.177	.785	.459
12. The corrections have been unclear.	10.874	2	5.437	4.157	.019*
13. The evaluation of the work is subjective.	3.241	2	1.620	.990	.376
14. It requires me to participate in my own evaluation (self-evaluation).	.005	2	.002	.004	.996

According to the results shown in Table 8, the items in which significant differences can be seen between the three academic years analysed are: *1.- It requires compulsory and active attendance; 8.- There is a disproportion between work/credits; and 12.- The corrections have been unclear.*

DISCUSSION

Regarding the first objective of the study on the analysis of the results of the applicability and difficulty of the TLP Good Practice, the students consider TLP to be useful, innovative, effective, sustainable and replicable. These results are the same as those found by other studies such as López-Pastor et al. (2020), López-Pastor et al. (2021) and Martínez-Mínguez (2016).

Regarding the category of competence acquisition, students consider that TLP and the F&SA system used help to acquire professional competences. In this sense, Gutiérrez et al. (2018) and Manrique et al. (2010) claim that the combination of good practices, such as TLP, and F&SA systems help the acquisition of professional competences of PTE students.

In the third category on evaluation and feedback, the results show that students value the support received from teachers more highly than from peers. Feedback is the most important part of the TLP development process, as it focuses on correcting and redirecting students' mistakes (Nicol et al., 2014; Panadero et al., 2017). In this sense, the results of the study do not coincide with those found by Bono and Núñez-Peña (2018), who state that peer feedback has an impact on learning and that, in addition, the feedback provided by the teacher converts it into internal feedback among peers because it has better results.

In short, the results show that students are quite satisfied with the TLP experience. This satisfaction with this good practice is also reflected in the study by Barba et al. (2010) and Martínez-Mínguez et al. (2019).

Regarding the ANOVA analysis to check if there are significant differences between academic years, it is worth noting that in the TLP data there are no significant differences between items, which reflects the consistency of the design of the good practice throughout the courses.

With regard to the second objective, concerning the analysis of the advantages and disadvantages of the F&SA system used in the same subject, it should be noted that the results are very positive. All the items analysed in relation to the advantages of the F&SA system obtained high averages, so the students consider that the F&SA system has several advantages, among them that learning is active and more is learnt, there is an interrelation between theory and practice and there is the possibility of correcting errors in the process and in the product. These advantages were already reflected in studies such as those by Nieva et al. (2021) and in a study a decade ago by Vallés et al. (2011).

As for the disadvantages that students point out about the F&SA system of the subject, they are: the requirement for compulsory and active attendance; the requirement for continuity and greater effort; the prior understanding of the assessment system; and the requirement to carry out self-assessments. These results are similar to those found by Romero et al. (2015). The rest of the items analysed obtained low and very low averages, so the students do not consider them to be disadvantages to be highlighted.

Over the academic years, the item concerning participation in the evaluation process has hardly changed in the mean over the academic years (mean of 4.50 or 4.51). In addition, it is worth noting that the item on the disproportion of work and credits has ambivalent results. The means range from 2.87 to 3.88, but the mode is 5 in the academic year 2019-2020, 1 in the academic year 2020-2021 and 3 in the academic year 2021-2022. In addition, the item Corrections have been unclear is the item with the lowest arithmetic means (between 1.56 and 2.40), which reflects the idea of feedback being applied throughout the course, both oral and written. Canabal and Margalef (2017) and Nicol et al. (2014) state that feedback helps to correct errors, because corrections should be as clear as possible, both by teachers and students.

As for the significant difference in the means after the one-factor ANOVA test, it should be noted that there are hardly any significant differences between the academic years. There are few significant differences in the advantages and disadvantages of the F&SA system implemented in the subject, with five and three differences respectively, which reflects the firmness of the design of the subject's assessment system over time.

CONCLUSIONS

In conclusion, the results found in the study are very positive, reaffirming the consistency of the subject design in terms of good practice and the F&SA system in place. The students consider that there are more advantages than disadvantages in the F&SA system used. The main limitation of the study is the small sample of students. As future research, it would be interesting to increase the number of years of the study in order to continue checking the consistency of its design.

This research may be of interest to PTE teachers who carry out TLP combined with F&SA systems or who wish to implement them in the classroom.

REFERENCES

- 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. *Movimiento*, 22(4), 1033–1048. Retrieved from <http://www.redalyc.org/articulo.oa?id=115349439002>
- Barrientos, E., López-Pastor, V.M., & Pérez-Brunicardi. (2019). ¿Por qué hago evaluación formativa y compartida y/o evaluación para el aprendizaje en EF? La influencia de la formación inicial y permanente del profesorado. *Retos*, 36, 37–43. Retrieved from <http://www.retos.org/>
- Bennett, R.E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles Policy and Practice*, 18(1), 1–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Canabal, C., & Margalef, L. (2017). La retroalimentación: La clave para una evaluación orientada al aprendizaje. *Procesos de Enseñanza-Aprendizaje. Estudios, Avances y Experiencia*, 21(2), 149–170. <https://doi.org/10.30827/profesorado.v21i2.10329>
- Cañadas, L. (2021). Contribution of formative assessment for developing teaching competences in teacher education. *European Journal of Teacher Education*, 44(3). <https://doi.org/10.1080/02619768.2021.1950684>
- Castejón, F.J., Santos-Pastor, M.L. & Cañadas, L. (2018). Desarrollo de competencias docentes en la formación inicial del profesorado de educación física. Relación con los instrumentos de evaluación. *Estudios Pedagógicos*, 44(2), 111–126. <http://dx.doi.org/10.4067/S0718-07052018000200111>
- Castejón-Oliva, F.J., Santos-Pastor, M.L., & Palacios Picos, A. (2015). Cuestionario sobre metodología y evaluación en formación inicial en educación física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, 15(58), 245–267. <https://doi.org/10.15366/rimcafd2015.58.004>
- Delgado, V., Ausín, V., Hortigüela-Alcalá, D., & Abella, V. (2016). Evaluación entre iguales: Una experiencia de evaluación compartida en Educación Superior. *Educadi*, 1(1), 9–24. <https://doi.org/10.7770/EDUCADI-V1N1-ART943>
- Fraile, A., López-Pastor, V.M., Castejón, F.J., & Romero, R. (2013). La evaluación formativa en docencia universitaria y el rendimiento académico del alumnado. *Revista Aula Abierta*, 41(2), 23–34. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=4239063>
- Gallardo, F., & Carter, B. (2016). La evaluación formativa y compartida durante el prácticum en la formación inicial del profesorado: análisis de un caso en Chile. *Retos*, 29(1), 258–263. Retrieved from <https://recyt.fecyt.es/index.php/retos/article/view/43550>
- Gallardo, F.J., López-Pastor, V.M., & Carter, B. (2020). Ventajas e Inconvenientes de la Evaluación Formativa, y su Influencia en la Autopercepción de Competencias en alumnado de Formación Inicial del Profesorado en Educación Física. *Retos*, 38, 417–424. Retrieved from <https://recyt.fecyt.es/index.php/retos/article/view/75540>
- Gutiérrez, C., Hortigüela-Alcalá, D., Peral, Z., & Pérez-Pueyo, A. (2018). Percepciones de Alumnos del Grado en Maestro en Educación Primaria con Mención en Educación Física sobre la Adquisición de Competencias. *Estudios Pedagógicos*, 44(2), 223–239. <http://dx.doi.org/10.4067/S0718-07052018000200223>
- Gutiérrez, C., Hortigüela-Alcalá, D., Peral, Z., & Pérez-Pueyo, A. (2018). Percepciones de Alumnos del Grado en Maestro en Educación Primaria con Mención en Educación Física sobre la Adquisición de Competencias. *Estudios Pedagógicos*, 44(2), 223–239. <http://dx.doi.org/10.4067/S0718-07052018000200223>
- Hamodi, C., & López, A.T. (2012). La evaluación formativa y compartida en la Formación Inicial del Profesorado desde la perspectiva del alumnado y de los egresados. *Psychology, Society, & Education*, 4(1), 103–116. Retrieved from <http://ojs.ual.es/ojs/index.php/psye/article/view/484>

- Hidalgo, M.E. (2021). Reflexiones acerca de la evaluación formativa en el contexto universitario. *Revista Internacional de Pedagogía e Innovación Educativa*, 1(1), 189–210. <https://doi.org/10.51660/ripie.v1i1.32>
- Hortigüela-Alcalá, D., González-Víllora, S., & Hernando-Garijo, A. (2021). Do we really assess learning in Physical Education? Teacher's perceptions at different educational stages. *Retos*, 42, 643–654. <https://doi.org/10.47197/retos.v42i0.88686>
- Hortigüela-Alcalá, D., Palacios, A., & López-Pastor, V.M. (2018). 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., & González-Calvo, G. (2019). Pero... ¿A qué nos Referimos Realmente con la Evaluación Formativa y Compartida?: Confusiones Habituales y Reflexiones Prácticas. *Revista Iberoamericana de Evaluación Educativa*, 12(1), 13–27. <https://doi.org/10.15366/riee2019.12.1.001>
- López-Pastor, V.M. (coord.) (2011). *Evaluación formativa y compartida en educación superior. Propuestas, técnicas, instrumentos y experiencias*. Narcea.
- López-Pastor, V.M., & Pérez-Pueyo, A. (Coords.) (2017). *Buenas prácticas docentes. Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas* (e-book). Universidad de León. Retrieved from <https://buleria.unileon.es/handle/10612/5999>
- López-Pastor, V.M., Fuentes, T., Pascual Arias, C., Molina Soria, M., & Fernández, C. (2021). La evaluación educativa en la formación inicial del profesorado. Ventajas y posibilidades de la evaluación formativa y compartida. In W. dos Santos, & R. Stieg (Org.), *Evaluación educativa en la formación de profesores* (cap.2, pp. 31–52). Appris.
- López-Pastor, V.M., Pascual Arias, C., Molina Soria, M., Fuentes, T., & Fernández, C. (2021). Cómo desarrollar sistemas de evaluación formativa y compartida en formación inicial del profesorado. Un ejemplo. In W. dos Santos, & R. Stieg (Org.), *Evaluación educativa en la formación de profesores* (cap.13, pp. 265–288). Appris.
- Lorente-Catalán, E., & Kirk, D. (2013). Alternative democratic assessment in PETE: an action-research study exploring risks, challenges and solutions. *Sport, Education and Society*, 18, 77–96. <https://doi.org/10.1080/13573322.2012.713859>
- Martínez, C. (2021). *Estrategias de retroalimentación y la evaluación formativa de las docentes del nivel de educación inicial de la provincia de Chíncha-2019* (Tesis Doctoral). Universidad Nacional Federico Villarreal
- Martínez-Mínguez, C., Moya, L., Nieva, C., & Cañabate, D. (2019). Percepciones de estudiantes y docentes: Evaluación formative en proyectos de aprendizaje tutorados. *Revista Iberoamericana de Evaluación Educativa*, 12(1), 59–84. <https://doi.org/10.15366/riee2019.12.1.004>
- Mastagli, M., Malani, D., Hainaut, J.P., Van Hoye, A., & Bolmont, B. (2020). Summative assessment versus formative assessment: An ecological study of physical education by analyzing stat-anxiety and shot-put performance among French high school student. *Journal of Physical Education and Sport*, 20(3), 2220–2229. <https://doi.org/10.7752/jpes.2020.s3298>
- Montero, I., & León, O.G. (2005). Sistema de clasificación del método en los informes de investigación en Psicología. *International Journal of Clinical and Health Psychology*, 5(1), 115–127. Retrieved from <http://www.redalyc.org/articulo.oa?id=33701007>
- Nicol, D. (2019). Reconceptualising feedback as an internal not an external process. *Italian Journal of Educational Research*, (numero speciale), pp. 71–84. <https://doi.org/10.7346/SIRD-1S2019-P71>
- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practice in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102–122. <http://dx.doi.org/10.1080/02602938.2013.795518>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144. <https://doi.org/10.1016/j.edurev.2013.01.002>

- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74–98.
<https://doi.org/10.1016/j.edurev.2017.08.004>
- Parra, S., & García-Martínez, I. (2021). Feedback and evaluative experience as decisive factors in student self-regulation. *Publicaciones*, 51(1), 303–316.
<https://doi.org/10.30827/publicaciones.v51i1.20738>
- Romero, R., Asún, S., & Chivite, M.T. (2016). La autoevaluación en expresión corporal en la formación inicial del profesorado de educación física: un ejemplo de buena práctica. *Retos*, 29, 236–241.
Retrieved from <http://www.redalyc.org/articulo.oa?id=345743464045>
- Saiz-Linares, A., & Susinos-Rada, T. (2018). Feedback and formative evaluation in a teachers reflexive practicum. *Avaliacao*, 10(30), 533–554. <https://doi.org/10.22347/2175-2753v10i30.1605>
- Winstone, N.E., & Boud, D. (2021). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*, 47(3), 656–667.
<https://doi.org/10.1080/03075079.2020.1779687>