Revista ELectrónica de Investigación y EValuación Educativa



e-Journal of Educational Research, Assessment and Evaluation

Differences of students and faculty on the difficulties to implement the formative assessment

Divergencias del alumnado y del profesorado universitario sobre las dificultades para aplicar la evaluación formativa

Romero-Martín, Rosario ⁽¹⁾; Castejón-Oliva, Francisco-Javier ⁽²⁾; López-Pastor, Víctor ⁽³⁾ (1) Universidad de Zaragoza. (2) Universidad Autónoma de Madrid. (3) Universidad de Valladolid

Resumen

La evaluación formativa supone un claro cambio en la forma de plantear la evaluación en la docencia universitaria. Su aplicación no está exenta de cierta controversia, y sobre todo de dificultades, que hacen que alumnado y profesorado tengan distintas perspectivas. El objetivo del presente estudio ha sido comprobar las divergencias del alumnado y del profesorado cuando se diseña y se ponen en práctica sistemas de evaluación formativa. Se ha utilizado un cuestionario para el alumnado (N=3013) y los informes elaborados por el profesorado (N=46) aplicados durante el curso 2012-13. Los resultados muestran que la evaluación formativa es bastante exigente para el alumnado, aunque también muestran una alta satisfacción del alumnado con la misma, particularmente con la calificación. Un inconveniente que refleja el alumnado parece ser la acumulación de trabajo al final del proceso, lo cual resulta paradójico y requiere de una investigación más profunda en el futuro. El profesorado también muestra una alta satisfacción con este tipo de evaluación, aunque tengan una mayor carga de trabajo. Se observan escasas discrepancias entre los dos colectivos (alumnado y profesorado). En el estudio se proponen recomendaciones de puesta en práctica en cuanto al diseño y la planificación para investigaciones futuras. Las limitaciones se observan en cuanto a la población participante en el estudio, con el fin de que puedan generalizarse los resultados.

Fecha de recepción 21 de Enero de 2014

Fecha de aprobación 20 de Abril de 2015

Fecha de publicación 23 de Abril de 2015

Palabras clave:

Evaluación Formativa; Evaluación participativa; Educación Superior; Percepción del estudiante universitario; Percepción del docente

Abstract

Formative Assessment shows a clear change in how to approach the evaluation of university teaching. The application of Formative Assessment is not without some controversy, and above all difficulties due to students and teachers having different perspectives. The objective of the present study was to check the differences of students and teachers when formative assessment systems are designed and put into practice. A questionnaire was used for students (N = 3013) and a report prepared by the teachers (N = 46). Both were taken from the academic year 2012-13. The results show that formative assessment is quite demanding for students, but also that students are highly satisfied with it, particularly with grades. A problem that the students reflect seems to be the accumulation of work at the end of the process, which is paradoxical and will requires further research in the future. Teachers also show a high satisfaction with this type of assessment, even if they have a greater workload. There are few differences between the two groups (students and teachers). The study proposes recommendations for implementation in terms of design and planning for future research. The limitations of the research can be seen in terms of participating in the study population, so the results can be generalized.

Reception Date 2014 January 21

Approval Date 2015 April 20

Publication Date: 2015 April 23

Keywords:

Formative Assessment; Participative assessment; Higher Education; Teachers Perception; University Students Perception

The implementation of formative assessment in higher education is not achieved without some difficulties, for both tutors and students. It is not just an issue of lack of knowledge, there is also resistance and a lack of tradition in its implementation (Capllonch et al., 2009; Pérez et al., 2008). This current study analyses the perceptions that students have about the use of formative assessment methods in university courses.

In this study it is assumed that the aim of formative assessment is to improve both the students' learning process and the tutors' (López-Pastor, teaching 2009). Previous studies have shown how the process focuses and re-focuses the learning and teaching process so that students achieve better learning and academic results (Torrance, 2012) and, furthermore, that it has a positive effect on the development of their skills and determination (Rué, 2009). Black & Wiliam (2009) and Yorke (2003) believe formative assessment should be approached delivered in a way that fully reflects its aims; for example, developing self-regulation and stimulating deep learning for students, which tutors acknowledge requires to independence of the students.

Some studies that endorse the use of formative assessment advocate that students should be directly involved in it (Biggs, 2005; Boud & Falchikov, 2007; Brown & Glasner, 2007; Knight, 2005) as they believe there is a relationship between the involvement of students and improvements in their learning (Bowden & Marton, 2012; Cano, 2012; Marton & Säljö, 1976). It appears that students' direct participation in formative assessment leads to improved learning and also to better academic results (Carrillo et al., 2009; Castejón, López-Pastor, Julián 2011; Romero-Martín, Zaragoza, López-Pastor & Castejón-Oliva, Aranda, 2014). Formative assessment requires appropriate feedback to be given so that students can improve their learning; feedback as part of formative assessment has been shown to deliver these improvements as long as it fulfils the criteria for effective feedback

Other previous studies have produced helpful findings with regard to implementing formative assessment in different subject areas or academic disciplines. In studies relating to medicine Krasner, Wimmers, Relan & Drake (2006) demonstrated that the use of formative assessment is a better predictor of a student's results than using summative assessment, that the strategies employed during formative assessment reduce tension and stress and make better use of learning time; something not achieved through summative assessment, which favours a good memory and being able recall information and the measurement of final results, in other words, "risking it all on one throw of the dice". Likewise, accounts of experimental studies on the use of formative assessment on social science courses indicate that its results are better than when traditional assessment methods are used (Santos, Martínez & López-Pastor, 2009).

In order to get students to be responsible for their own learning one valuable strategy is to implement formative assessment, selfassessment and peer assessment. Fallows y (2001)Chandramohan suggest that the introduction of self- and peer assessment delivers advantages that more than justify their inclusion in higher education courses. Studies that students generally positively, particularly those in their final year. Nevertheless, students do not participate with the same level of enthusiasm and differences between the approach of students and tutors can become apparent, either through students' reluctance to assess their peers or for the workload that these assessment methods generate (Gijbels & Dochy, 2006).

The use of formative assessment generally leads to improvements in students' academic performance, particularly when appropriate criteria have been established for implementation (Sadler, 2005). Several studies undertaken concerning initial teacher training identify significant improvements in students' marks (Castejón et al., 2011; Julián, Zaragoza, Castejón & López-Pastor, 2010; López-Pastor, 2008; López-Pastor, Manrique & Vallés, 2011), although they also acknowledge that students can demonstrate a certain resistance when participating in this type of assessment and respond negatively to what the tutor proposes. It therefore appears essential to understand students' viewpoints to resolve this apparent paradox and to ensure learning continues, especially when their viewpoints conflict with those of their tutors.

Previous studies investigated have students' perceptions regarding the use of formative and shared assessment methods in higher education and their influence on variables such as: the perception of learning, participation and contribution in the process and the workload that it generates. One example is the work of Hamodi & López-Pastor (2012), which uses group discussion as a tool for obtaining data. The participants in this study were students and graduates of initial teacher education. The authors found that students do not always understand tutors when they introduce assessment, especially when the discourse concerns the need for continuous assessment when what they actually do is to "risk all" the module on a single final exam, possibly linked to some work that students do separately and which they have to pass in order to achieve an average mark. This suggests that differences are apparent not only in terms of students' perceptions of assessment but also due to the lack of training of tutors, as alluded to in the past (Fernández, 1989; Tejedor, 1998) and which remains a critical issue in the initial training and development of tutors (De la Calle, 2004; Tonucci, 2010; Trillo, 2005). Other studies have produced similar findings (Boussada & De Ketele, 2008). However, some studies do indicate a connection between what they have learned about formative assessment during their initial education and the way they implement it during their professional careers (Hamodi & López-Pastor, 2012).

On the other hand, tutors do use assessment methods to determine the learning of their students that are, at times, contrary to institutional guidelines (models) or their personal practices (pedagogic training). Consequently, Palacios & López-Pastor (2013) suggest there are three types of tutor: (a) innovator (25%), with positive attitudes towards continuous and formative assessment; (b) traditional (26%), with negative attitudes towards continuous and formative assessment; and (c) eclectic (49%), those whose position is halfway between these previous attitudes. Tutors that are "innovators", are more predisposed towards having their students actively participate in the assessment process whilst "eclectic" tutors are less predisposed and usually combine innovative methods with a final exam, whilst "traditional" tutors prefer to use a final exam and/or another assignment as a tool for assessing and marking students.

which combine Studies the two (tutors' perspectives and students') consistently indicate clear differences between them. The papers reviewed seem to suggest that students do consider formative assessment be important and tutors agree differences are apparent and which are revealed in a resistance on the part of both groups (Struyven, Dochy & Janssens, 2005). For example, Gutiérrez, Pérez, Pérez & Palacios (2011) show that tutors are more open to proposing alternative assessment methods but acknowledge that students do not always recognise them; whilst students perceive that their assessment methods are "the same as always". Tutors also mention, though, that students show a lack of commitment when, for example, the focus is on achieving deeper learning (Gijbels, Segers & Struyf, 2008; Lizzio, Wilson & Simons, 2002), and even to the institution itself (Martínez, Castejón & Santos, 2012).

With these issues in mind it appears there is a need to investigate why it is that students do not fully engage with an assessment process that has demonstrable benefits for their learning. The aim of this study, therefore, is to explore the different reactions of students and

tutors when formative assessment processes are designed and implemented and to understand both the resistance they show towards them and the drawbacks they feel they have.

Methods

Participants

The sample in this study comprises 3,013 students enrolled on a range of degree courses with the majority taking Initial Teacher Education courses (ITE). All participants were asked to rate the formative assessment they experienced on appropriate modules during the 2012-13 academic year. The sample is intentional as the students and tutors involved were specifically chosen for this study. The student participants came from the areas of Teacher Education (32 universities), Physical Activity and Sport Science (6 universities) and other disciplines (8 universities). The students came from every year, from first to final year, of their respective degree courses. Responses were received from 100% of the students that regularly attended their course lectures.

All 46 tutors who participated belong to National Network for Formative Assessment in Higher Education (Spain). This means they are all obliged to implement formative assessment in their modules and present a structured report on their personal experiences of using it with their students during the 2012-2013 academic year. The students were told at the beginning of the academic year that formative assessment would be used in the module and they were given information with practical examples so that they could understand it and share their opinions on the concept.

Data collection instruments

Two instruments were used to obtain data: (a) the individual reports of the member of the National Network for Formative Assessment in Higher Education and (b) the questionnaire applied to the students for them to rate the module.

The individual member reports from the National Network for Formative Assessment in Higher Education. This is a highly structured individual and personal report on the results of implementing formative and assessment within a semester or year-long module. Data is collected during the module and the report produced at the end. This instrument has been previously validated and applied in other published research which means there are precedents that guarantee the validity of these reports (Buscà, Pintor, Martínez, & Peire, 2010; Manrique, Vallés & Gea, 2012; Vallés, Ureña & Ruiz, 2011). The individual reports contain the following Information sections: (1) regarding (2) Correlation between context: assessment and the learning; (3) Learning and assessment strategies; (4) Implementation of assessment instruments and the focus of the learning; (5) Ethical criteria and the type of decisions made assessment; in Development of the assessment strategy using ICT; (7) Advantages identified in the assessment strategy used; (8) Disadvantages identified in the assessment strategy used; (9) Academic performance and its measurement; (10) Workloads for students and tutors; and (11) Conclusions.

In section 7 the tutors express their personal degree of satisfaction with the implementation of formative assessment and in section 8 they consider the experience from the students' perspective, in both sections using a five point Likert scale: *not at all, little, some, quite a lot* and *a lot*. Furthermore, in both sections they refer to the results of an anonymous questionnaire their students have completed about the same experience and which is described below.

Questionnaire on methodology and assessment in Physical Education Initial Teacher Education. This is an anonymous questionnaire which students use to rate a module. It was validated by Castejón, Santos & Palacios (at press) achieving a Cronbach's Alpha Coefficient score of 0.84. The questionnaire was subjected to a three part validation process for content, understanding

and reliability: (a) validity of the content by experts in the subject matter; (b) a pilot study aimed at defining the questions; and (c) application of the questionnaire to 892 students on the Initial Teacher Education degrees of a range of universities. A Factor Analysis of Principal Components was undertaken with the final version of the questionnaire and acceptable values were obtained for both the KMO measure, 0.807 and Bartlett's Test of Sphericity (p>.00). The values obtained from the covariance matrix produced acceptable adjustments for RMSEA = 0.078, and GFI = 0.88, and for all other model adjustment indicators.

Students indicate their level of agreement with the statements provided using a five point Likert scale (not at all, little, some, quite a lot and a lot) in relation to the advantages and disadvantages of the assessment process used. The complementarity of the two instruments enables the students' perspective to be identified and compared with that of the tutors'.

Scope and Variables

The scope and variables refer to the aspects considered as advantages or disadvantages of implementing formative assessment identified through the tutors' reports from previous years, from research meetings and the specialised bibliography. This study focuses on three (see each item in Table 1).

- 1. Disadvantages Students. The students' level of agreement with the statements on formative assessment regarded as disadvantages (14 items).
- 2. Disadvantages Tutors. Tutors' opinions: (1) Personal perception of the workload generated by the process; (2) Perception of the resistance shown by the students to formative assessment (2 items).
- 3. Advantages Tutors. Tutors' opinions: Personal level of satisfaction with the process (1 item).

Table 1. Scope, items and variables covered by the aims of this study

| Scope | Variab | espe, rems and variables especially are aims of this study |
|---------------------|------------|--|
| Disadvantages - | DS1 | It requires compulsory and active attendance |
| Students | DS2 | It has a work dynamic that is not widely understood and people are unfamiliar with |
| | DS3 | It requires continuity |
| | DS4 | It needs to be explained beforehand |
| | DS5 | It requires a greater effort |
| | DS6 | Group working causes difficulties |
| | DS7 DS8 | A lot of work accumulates at the module end |
| | | There is an imbalance between the workload and the credits awarded |
| | DS9 | The marking process is more complex and sometimes unclear |
| | DS10 | It creates insecurity and uncertainty and doubts about what is required |
| | DS11 | It is unfair compared to other assessment methods |
| | DS12 | The corrections were not clear |
| | DS13 | The marking of the work is subjective |
| | DS14 | It requires students to participate in their own assessment (self-assessment) |
| Disadvantages - | DT1 | It creates an excessive workload for tutors |
| Tutors | DT2 | Students show strong resistance to formative and continuous assessment methods |
| Advantages - Tutors | AT1 | The satisfaction of the tutors with the assessment methods they use |

Procedure

An initial model for an individual report was designed by a working group from the Research & Development Plus Innovation Project on formative assessment in initial teacher education and was used for a pilot study. The model was subsequently modified and changed and the final version was used for this current study. The definitive version of the report (explained in the "Instruments" section) was completed by each tutor after finishing the module in which they had included formative The deadlines for completion assessment. across the Network for Formative Assessment are the end of the first or second semester, depending on the duration of the module. The students complete their anonymous questionnaires at the end of each semester. On average students take 30 to 40 minutes to complete it. The data was sent to the coordinator who, together with the data analysis team, input all the responses into the appropriate software package.

Data processing

To process the data five points on the ordinal scale with numeric values were

identified; the score 0 corresponds to the option not at all; 1 to little, 2 to some; 3 to quite a lot and4 to a lot). The analytical procedures used by Gutiérrez, Pérez & Pérez (2013) for a similar study involving students, graduates and tutors were used as a reference point together with Gutiérrez et al. (2011) for the treatment of ordinal variables. Descriptive comparative reports wereprepared and covering the opinions of the students and the tutors specifically: (1) Descriptive of the variables for both tutors and students (Averages and Standard Deviation); (2) Comparative between the items for the tutors (Spearman's Rho test); (3) Comparative between the variables for the tutors and students (Kruskal-Wallis Test). Both Excel_2007 and SPSS_v19 programmes were used, with a level of significance of p

Results

Descriptive study of the students' perceptions

Figure 1 presents the results of the level of agreement of the students with 14 statements regarded as disadvantages of using formative assessment in higher education.

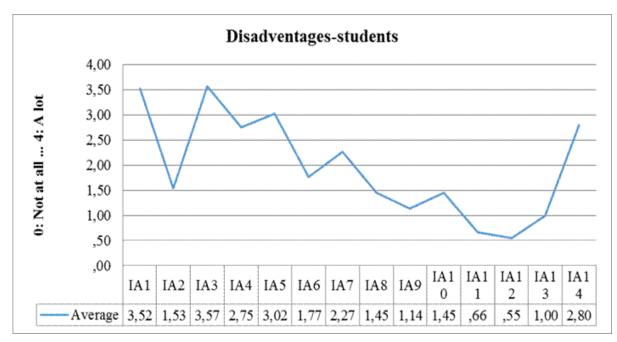


Figure 1. Level of agreement or disagreement of students with the potential disadvantages of formative assessment.

As can be seen, for eight of the 14 items the scores are below 2 which corresponds to some; for three items they are between some and quite a lot and for a further three items they are between quite a lot and a lot. For more than half of the items the scores given by the students to the disadvantages are low and for only three, slightly more than 20%, they give high scores, producing an overall average of M = 1.87, which is below some. To help with understanding the data they have been arranged in three sections: (a) those relating to aspects of the process and organisation; (b) those relating to the requirements of the assessment strategy; and (c) those relating to marks.

Process and organisation (implementation). Table 2 presents the data relating to the possible disadvantages in implementation (procedures, organisation, etc.) of formative assessment in higher education (items 2, 4, 6, 7, 8 and 10). They show four scores between little and some; the two nearest to little refer to the imbalance between the workload and the credits given for the module (M = 1.45) and to whether it produces greater certainty or uncertainty (M = 1.45). The two closest to some relate to whether the assessment method has a work dynamic that is not widely understood (M = 1.53) and to the difficulty of group working (M = 1.77).

Table 2. Potential disadvantages in the implementation of formative assessment in higher education

| Item | Variable | Avge. | Std Dev. |
|------|--|-------|----------|
| DS2 | It has a work dynamic that is not widely understood and people are unfamiliar with | 1.53 | 1.162 |
| DS4 | It needs to be explained beforehand | 2.75 | 0.651 |
| DS6 | Group working causes difficulties | 1.77 | 1.008 |
| DS7 | A lot of work accumulates at the module end | 2.27 | 1.042 |
| DS8 | There is an imbalance between the workload and the credits awarded | 1.45 | 0.926 |
| DS10 | It creates insecurity and uncertainty and doubts about what is required | 1.45 | 0.901 |

Among the items that show a positive level of agreement with potential disadvantages is item DS7 which relates to the possible accumulation of work at the end of the module (M = 2.27). Similarly the students showed they were generally in agreement with item DS4, that it is important to understand formative assessment processes in advance (IA4, M = 2.75).

Level of requirement. Figure 1 shows the highest scores the students attribute to all of

the statements relating to disadvantages are given to the four that include a reference to specific requirements (items 1, 3, 5 and 14; in Table 3). The highest level of agreement was that formative assessment methods require continuity (M = 3.57), compulsory and active attendance (M = 3.52), a greater effort (M = 3.02) and involvement in their own assessment, that is, self-assessment (M = 2.809).

Table 3. Level of agreement with the requirements demanded by formative assessment.

| Item | Variable | Average | Std. Dev. |
|------|---|---------|-----------|
| DS1 | It requires compulsory and active attendance | 3.52 | .762 |
| DS3 | It requires continuity | 3.57 | .587 |
| DS5 | It requires greater effort | 3.02 | .590 |
| DS14 | It requires students to participate in their own assessment | 2.80 | 1.112 |

Marks. Table 4 shows how students consider of low importance the possible disadvantages for marks that formative assessment methods might have (items 9, 11, 12 and 13). All are given very low scores,

especially in terms of the potential lack of fairness compared with other assessment methods (M = .66) and in relation to the lack of clarity of corrections (M = .55).

Table 4. Potential disadvantages relating to the marking system.

| Item | Variable | Average | Std. Dev. |
|------|---|---------|-----------|
| DS9 | The marking process is more complex and sometimes unclear | 1.14 | .930 |
| DS11 | It is unfair compared to other assessment methods | 0.66 | .861 |
| DS12 | The corrections were not clear | 0.55 | .820 |
| DS13 | The marking of the work is subjective | 1 | .964 |

The tutors' perspectives

In terms of the three variables relating to the tutors it can be seen from Table 5 that tutors who use formative assessment in their modules show a high degree of satisfaction (M = 3.31).

With regard to the disadvantages mentioned by the tutors the data show that the majority of them feel that using formative assessment increases their workload (M = 3.02); however, they do not consider that the students are strongly resistant to formative assessment methods (M = 1.18).

Table 5. Tutors' opinions: level of satisfaction and potential disadvantages

| Item | Variable | Average | Std. Dev. |
|------|--|---------|-----------|
| AT1 | Satisfaction of the tutors with the assessment methods they use | 3.31 | .569 |
| DT1 | Creates an excessive workload for tutors | 3.02 | .79018 |
| DT2 | Students show strong resistance to formative and continuous assessment methods | 1.18 | .95451 |

Comparison between the tutors' opinions

Table 6 indicates the tutors' level of satisfaction with the formative assessment

method they use in relation to whether they feel the workload it creates to be excessive (AT1 vs. DT1)

Table 6. Comparison between the tutors' level of satisfaction and the workload generated

| Comparison AT1 vs. | Avge. AT1 | Avge. DT1 | Coefficient of correlation | Sig. (bilateral) |
|---------------------|--------------|-----------|----------------------------|------------------|
| DT1_ Spearman's Rho | 3.31 | 3.02 | -0.067 | 0.688 |

The average value for both variables is high (see Table 6), being above 3, some/quite a lot. The value of the coefficient of correlation, at close to zero and negative, indicates a very weak inverse correlation, but there is no significance, as shown by p = .688, meaning that these variables have no

correlation and that tutors hold pairs of heterogeneous opinions.

Comparison between tutors' and students' opinions

Table 7 compares the tutors' perception of students' resistance to formative assessment to the students' views and also attempts to

identify the issues that could cause this resistance.

According to the tutors the students would show little opposition to formative assessment, as evidenced by their score for this variable (DT2) M = 1.18, whilst the overall

average score for the students' opinion was greater (M = 2). To investigate this further the tutors' variable (DT2) is compared with each of the students' variables (DS, DS2..., DS14), using the Kruskal Wallis test, with the results expressed in Table 7.

Table 7. Comparison between the tutors' perceptions and the opinions of the students regarding resistance to formative assessment

| | | | Summary of the Kruskal Wallis hypothesis test - significance .05 | |
|--------------|-------------|---------------|--|----------|
| Comparison | Avge. tutor | Avge. student | Sig. | Decision |
| DT2 vs. DS1 | 1.18 | 3.52 | .055 | accept |
| DT2 vs. DS2 | 1.18 | 1.53 | .945 | accept |
| DT2 vs. DS3 | 1.18 | 3.57 | .370 | accept |
| DT2 vs. DS4 | 1.18 | 2.75 | .231 | accept |
| DT2 vs. DS5 | 1.18 | 3.02 | .065 | accept |
| DT2 vs. DS6 | 1.18 | 1.77 | .356 | accept |
| DT2 vs. DS7 | 1.18 | 2.27 | .031* | reject |
| DT2 vs. DS8 | 1.18 | 1.45 | .161 | accept |
| DT2 vs. DS9 | 1.18 | 1.14 | .461 | accept |
| DT2 vs. DS10 | 1.18 | 1.45 | .779 | accept |
| DT2 vs. DS11 | 1.18 | 0.66 | .019* | reject |
| DT2 vs. DS12 | 1.18 | 0.55 | .018* | reject |
| DT2 vs. DS13 | 1.18 | 1.00 | .061 | accept |
| DT2 vs. DS14 | 1.18 | 2.80 | .113 | accept |

It can be seen that in 11 of the 14 comparisons the significance level is greater than p = .05 so in those 11 cases the hypothesis of equality can be accepted, which means there are no significant differences between the tutors' perceptions and the students' opinions. However, in the remaining three cases marked with an asterisk, DS7 (A lot of work accumulates at the module end; M = 2.27); DS11 (It is unfair in comparison to other assessment methods; M = .66), DS12 (Corrections were not very clear; M = .55), discrepancies between the tutors and the students are evident.

To understand more about these discrepancies a comparison is made between the average scores of the students and the tutors which indicates that the students placed more importance than their tutors on the issue

that a lot of work accumulates at the module end (DT2 vs DS7). In contrast the students gave less importance to the issues of clarity or the unfairness in the marking (DT2 vs. DS11 & DS12).

Discussion

The students' perceptions

In previous studies on students' perceptions of possible disadvantages in formative assessment some authors highlight that students demonstrate some resistance when participating in this form of assessment (Castejón et al., 2011; Julián et al., 2010; López-Pastor, 2008; López-Pastor et al., 2011; Pérez et al., 2008; Capllonch et al., 2009); but in this study the students did not present a high level of resistance in general, but rather only in specific aspects, reflecting other published

research which indicates that the resistance appears at various points in the process, from the design and implementation (López-Pastor, 2009; Sadler, 2005), to feedback (Carless, 2006).

Procedures and organisation. In terms of procedures and organisation, four issues, which priori might appear to disadvantages were not highlighted by the students; they found no major problems with group working, they had few uncertainties or doubts about what they had to do and they did not feel there was any major disproportion between the workload and the credits, that all reflects other studies (Julián et al., 2010; López et al., 2011; Romero-Martín et al., 2014), which suggest that using formative assessment does not create an excessive workload for students, but that it falls within the parameters for the number of study hours per ECTS credits for each module.

Students did not highlight as a major disadvantage the lack of familiarity with this form of assessment, which coincides with the work of Gutiérrez et al. (2011); Palacios & López-Pastor (2013) and Palacios, López-Pastor & Barba (2013), concerning the gradual increase in the percentage of tutors in Spanish universities that use formative and continuous assessment strategies.

With regard to the distribution of the work throughout the academic year the results suggest that at least a small number of tutors that used formative assessment did not spread appropriately throughout both semesters. This creates a greater accumulation of tasks towards the end of the year, precisely at the time when students have to spend more time preparing for their traditional final exams. This is evidenced in Hamodi & López (2012), which identified and analysed students' negative opinions on this tradition in Spanish universities and also in Julián et al. (2010), which demonstrated that if fixed deadlines for each assignment are not established in advance for the whole year, it leads to an excessive accumulation of work at the end of the year and to increased failure or withdrawals. All of this reinforces the idea that it is vital to set deadlines for undertaking and submitting the majority of assignments in the first three months of the academic year to avoid the problem of an excessive workload at the end of year.

The students considered it important that they understood the assessment method in advance, as identified by Castejón et al. (2011), who place significant importance on making the whole process very clear, thus avoiding being over confident in it being successfully received, as cautioned by Buscà et al. (2010).

It is not possible to conclude therefore whether the factors relating to the organisation of assessment systems produce disadvantages that are unsurmountable, although some items provided important data on where improvements can be made the in implementation of formative assessment.

Level of requirement. The students confirm that formative assessment requires a greater effort from them than other assessment formats and that it requires continuity of work and class attendance and that their attendance needs to be active; meaning that the student's involvement cannot remain purely superficial. This, in turn, leads to processes of selfregulation being required (Black & Wiliam, & Yorke, 2003), which impacts 2009 positively on students' involvement, reducing student failure and withdrawals, enhancing their learning (López-Pastor, Pintor, Muros & Webb, 2013) and their overall academic performance (Biggs, 2005; Boud & Falchikov, 2007; Bowden & Marton, 2012; Brown & Glasner, 2007; Cano, 2012; Carrillo et al., 2009; Castejón et al., 2011; Knight, 2005; Marton y Säljö, 1976; Romero-Martín et al., 2014).

Marking. The results contradict the prevailing opinions about this form of assessment. In spite of the volume and range of learning activities normally undertaken when formative assessment is used the students in this study did not believe the marking process was more complex or unclear. On the contrary, they indicated that the marking of their work had not been as

subjective and they highlighted as a positive issue the general fairness of the process and the clarity of the corrections. This tends to suggest that the tutors had designed and implemented effective teaching strategies for using formative assessment, in accordance with the recommendations in Castejón et al. (2011)

These findings are particularly interesting given the well-known obsession with marks; both on the part of students and also the tutors who need to ensure they reflect the true effect of the teaching and learning process for each student.Several studies about the use of formative assessment in the initial training of physical education teachers actually found significant improvements in students' marks (Castejón et al., 2011; López-Pastor, 2008; López-Pastor et al., 2011), which in this study is reflected by the students' positive perception marking process in formative of the assessment.

Tutors' opinions

In this study the tutors indicated a high satisfaction of with formative assessment; but it should be remembered that tutors participating had undertaken voluntarily to implement formative assessment strategies and therefore reflected the profile of "innovative tutors" according Palacios & López-Pastor terminology of (2013) and Palacios et al. (2013), who define it as tutors with positive attitudes towards continuous and formative assessment, meaning they had overcome the resistance as mentioned by Pérez et al. (2008) and Capllonch et al. (2009); therefore the high level of satisfaction they recorded is as might be expected.

The tutors also felt their workload was higher, reflecting other studies that suggest the main disadvantage of formative assessment is the additional time and commitment required of them (Bennet, 2011; Gibbs & Simpson, 2004-05).

Comparing these two variables shows a slight tendency towards tutors being generally satisfied with implementing formative assessment in spite of the increase in workload

it creates. This is possibly because they are aware the workload actually falls within the parameters set for ECTS credits for their modules, as shown by research published by Julián et al. (2010) and López-Pastor et al. (2013), which use authentic data on the actual workload to demonstrate that it is perfectly viable. The correlation, though, is not significant, which means it is not possible to fully confirm this assertion.

Comparison between the tutors' perceptions – and the students' opinions

Although some literature shows tutors perceived some resistance from students towards formative assessment (Capllonch et al., 2009; Pérez et al., 2008; Struyven et al., 2005), the tutors in this study perceived little resistance, possibly because their day to day observation suggested a good general level of acceptance, which the students went on to confirm. The results indicate that the students' opinions matched the tutors' perceptions to a high degree.

Of the three instances where discrepancies are found, in two of them the students score those disadvantages even lower than the tutors, particularly for "It is unfair in comparison with assessment methods" and corrections have not been very clear". It is important to recognise that this is an unexpected result because marks often cause disagreements between students and tutors, but if formative assessment demands that students should become critical and independent learners (Torrance, 2012), it is reasonable to expect that their opinions were given based on that situation.

In the third instance where there was a discrepancy between tutors and students, in relation to the accumulation of work at the end of a semester, the students' scores were higher than the tutors'. This is particularly relevant when considering that one of the basic features of formative assessment is continuity, in as far as it enables constant feedback to be given which constantly refocuses students' learning and continually reduces the gap between what they do know and what they should know

(Taras, 2009; Torrance, 2012); in summary, it maximises the students' use of information to enhance their learning which, in turn, increases their motivation and performance, as shown in studies that focus on this aspect (for example, Burke, 2009; Carless, 2006; Carless et al., 2010).

These results enable tutors to modify their perceptions and suggest ways in which the formative assessment processes they implement can be improved.

Conclusions

Scientific research into the effects of formative assessment appears to be still inconclusive. So far, studies have focused on comparing advantages and disadvantages or providing additional evidence to support other studies. This present study, reflecting the approach in Dunn & Mulvenon (2009), has produced data based on a concrete experience, providing valuable information so that tutors can implement formative assessment knowing that they need to do so carefully and with arguments ready to overcome resistance. Just as this study intended, it has proved that there are differences in the opinions of tutors and students regarding formative assessment. Students tend to show resistance to formative assessment, although in this study it was less than expected in relation to its implementation and the increased workload, possibly due to them having experienced this assessment strategy in various modules and therefore being more accustomed to it. However, the results also indicate that it is beneficial if tutors spread the workload more carefully over the whole year, to avoid overloading the final weeks of the semester when continuous assessment can clash with end summative formal of year assessments in other modules. Students' perception of the workload seems to improve when deadlines are set that enable them to cope with the requirements of formative assessment. It also reduces the rates of withdrawal and failure that are seen when deadlines are too closely bunched

Another of the study's aims is also proven; that when students are involved in the design and implementation of assessments they become better disposed towards it. Despite them still considering it a method that demands constant attendance and dedication. this does not seem to make them perceive it overall in a negative way. experience of formative assessment actually improves their self-regulation in terms of assessment and learning. Students also have a positive view on the relationship between formative assessment methods and their marks; they do not see it as a less fair method or indicate any other concerns about it. They are particularly positive about the relationship between the level of involvement, the range of learning tasks incorporated and the marks they receive.

Meanwhile, the tutors believe formative assessment implies a high level of commitment but also a high degree of satisfaction because of the advantages it has for students. Moreover, it appears that the level of commitment required is not inconsistent with the ECTS parameters, although the data here are less clear than in other published studies.

Comparing the opinions of tutors and students in our study shows a high degree of agreement between them on around 80% of the items and significant differences on only the remaining 20%. These concern: it is unfair in comparison with other assessment methods, the corrections were not very clear and the workload can accumulate at the end of the semester. But these issues can be easily resolved, as shown throughout this paper, through more careful planning on the part of tutors and greater involvement on the part of students.

Among the limitations of this current study is that the participating tutors have been using formative assessment for some time and therefore their responses should be contrasted with those of tutors that have not implemented it or those that have a variety of degrees of experience in using formative assessment. Likewise, a comparison should be made

between students that have some or no previous experience of formative assessment. It would also be valuable to investigate other relevant issues such as help for tutors when planning the tasks and assignments that students have to submit and with the type of feedback they should give in order to ensure students develop greater self-regulation of their learning and development.

References

- Bennet, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*, 5-25. DOI:10.1080/0969594X.2010.513678
- Biggs, J. (2005). Calidad del aprendizaje universitario. Madrid: Narcea.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*, 5-31. DOI: 10.1007/s11092-008-9068-5
- Boud, D. & Falchikov, N. (2007). Rethinking Assessment in Higher Education. Learning for the long term. Londres: Routledge
- Boussada, H. & De Ketele, J-M. (2008). L'évaluation de la qualité de la formation et du système d'évaluation universitaire: Le point de vue des diplômés. *Avaliação*, *13*(1), 39-61. DOI:10.1590/S1414-40772008000100003
- Bowden, J. & Marton, F. (2012). *La Universidad. Un espacio para el aprendizaje*. Madrid: Narcea.
- Brown, S. & Glasner, A. (Eds.) (2007). Evaluar la Universidad. Problemas y nuevos enfoques (2ª ed.). Madrid: Narcea.
- Burke, D. (2009). Strategies for using feedback students bring to higher education. *Assessment & Evaluation in Higher Education*, 34(1), 41-50. DOI: 10.1080/02602930801895711
- Buscà, F., Pintor, P., Martínez-Mínguez, L. & Peire, T. (2010). Sistemas y procedimientos de Evaluación Formativa en docencia universitaria: Resultados de 34 casos aplicados durante el curso académico 2007-2008. Estudios sobre Educación, 18, 255-276. Retrieved from http://dspace.unav.es/dspace/bitstream/10171/9829/2/ESE 18 11.pdf

- Cano, E. (Ed.) (2012). Aprobar o aprender. Estrategias de evaluación en la sociedad red. Barcelona: Laboratori de Mitjans Interactius.
- Capllonch, M., Ureña, N., Ruiz, E., González, N., Pérez, A., López, V. M & Tabernero, B. (2009). Guía de problemas y soluciones para el desarrollo de sistemas de evaluación formativa en la universidad. En V. M. López Pastor, (coord.), La Evaluación Formativa Universitaria: Compartida enDocencia propuestas, técnicas, instrumentos experiencias (pp. 217-252). Madrid: Narcea.
- Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education,* 31(2), 219-33. DOI: 10.1080/03075070600572132
- Carless, D., Salter, D., Yang, M. & Lam, J. (2010). Developing sustainable feedback practices. *Studies in Higher Education*, *36*(4), 395-407. DOI: 10.1080/03075071003642449
- Carrillo, M.T., Baillès, E., Caseras, X., Martínez, A., Ortet, G. & Pérez, J. (2009). Formative assessment and academic achievement in pregraduate students of health sciences. *Advances in Health Sciences Education*, *14*, 61-67. DOI: 10.1007/s10459-007-9086-y
- Castejón, F. J., López-Pastor, M. L.; Julián, J. & Zaragoza, J. (2011). Evaluación formativa y rendimiento académico en la Formación inicial del profesorado de Educación Física. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte, 11(42), 328-346. Retrieved from http://cdeporte.rediris.es/revista/revista42/artevaluacion163.htm
- Castejón, F. J., Santos, M.L. & Palacios, A. (En prensa). Cuestionario sobre metodología y evaluación en formación inicial en educación física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*. Retrieved from http://cdeporte.rediris.es/revista/inpress/artescal a566.pdf.
- De la Calle, M.J. (2004). El reto de ser profesor en el contexto de la convergencia europea. La formación pedagógica como necesidad. *Revista Universitaria de Formación del Profesorado, 18*(3), 251-258. Retrieved from http://www.redalyc.org/articulo.oa?id=27418316
- Dunn, K. E. & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical*

- Assessment, Research & Evaluation, 14(7), 1-11.
- Falchikov, N. y Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322. DOI: 10.3102/00346543070003287
- Fallows, S. & Chandramohan, B. (2001). Multiple Approaches to Assessment: reflections on use of tutor, peer and self-assessment. *Teaching in Higher Education*, 6(2), 229-246. DOI: 10.1080/13562510120045212
- Fernández, M. (1989). Así enseña nuestra universidad: Hacia la construcción crítica de una Didáctica universitaria. Madrid: UCM.
- Gibbs, G. & Simpson, C. (2004-05). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-31.
- Gijbels, D. & Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference? *Educational Studies*, 32(4), 399-409. DOI: 10.1080/03055690600850354
- Gijbels, D., Segers, M. & Struyf, E. (2008). Constructivist learning environments and the (im)possibility to change students' perceptions of assessment demands and approaches to learning. *Instructional Science*, 36, 431-443. DOI: 10.1007/s11251-008-9064-7
- Gutiérrez-García, C., Pérez-Pueyo, A. & Pérez-Gutiérrez, M. (2013). Percepciones de profesores, alumnos y egresados sobre los sistemas de evaluación en estudios universitarios de formación del profesorado de Educación Física. Ágora para la Educación Física y el Deporte, 15(2), 130-151. Retrieved from http://agora-revista.blogs.uva.es/files/2013/11/agora-15-2d-gutierrez-et-al.pdf
- Gutiérrez, C., Pérez, Á., Pérez, M. & Palacios, 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. <u>DOI:</u> 10.1174/113564011798392451
- Hamodi, C. & López-Pastor, A. T. (2012). La evaluación formativa y compartida en la Formación Inicialdel Profesorado desde la perspectiva del alumnado y de los egresados.

- Psychology, Society, & Education, 4(1), 103-116
- Knight, P. T. (2005). *El profesorado de Educación Superior*. Madrid: Narcea.
- Krasner, S., Wimmers, P. F., Relan, A. & Drake, T. A. (2006). Differential Effects of Two Types of Formative Assessment in Predicting Performance of First-year Medical Students. *Advances in Health Sciences Education*, 11, 155-171. DOI: 10.1007/s10459-005-5290-9
- Julián, J., Zaragoza, J., Castejón, F. J. & López-Pastor, V. M. (2010). Carga de trabajo en diferentes asignaturas que experimentan el sistema ECTS. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte, 10(38), 218-233. Retrieved from http://cdeporte.rediris.es/revista/revista38/artcarga151.htm
- Lizzio, A., Wilson, K. & Simons, R. (2002). University Students' Perceptions of the Learning Environment and Academic Outcomes: implications for theory and practice. *Studies in HigherEducation*, 27(1), 27-52. DOI: 10.1080/0307507012009935 9
- López-Pastor, V. M. (2008). Desarrollando sistemas de evaluación formativa y compartida en la docencia universitaria. Análisis de resultados de su puesta en práctica en la formación inicial del profesorado. *European Journal of Teacher Education*, 31(3), 293-311. DOI: 10.1080/02619760802208452
- López-Pastor, V. M. (Ed.) (2009). Evaluación Formativa y Compartida en Educación Superior. Madrid: Narcea.
- López-Pastor, V. M., Manrique, J.C. & Vallés, C. (2011). Evaluación y calificación en los nuevos estudios de Grado. Especial incidencia en la Formación Inicial del Profesorado. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 39(14.4), 1-20. Retrieved from http://www.aufop.com/aufop/revistas/arta/digital/165/1675
- López-Pastor, V. M., Pintor, P., Muros, B. & Webb, G. (2013). Formative assessment strategies and their effects 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. DOI: 10-1080/0309877X.2011.644780

- Manrique, J. C., Vallés, C. & Gea, J. M. (2012). Resultados generales de la puesta en práctica de 29 casos sobre el desarrollo de sistemas de Evaluación formativa en docencia universitaria. *Psychology, Society & Education, 4(1),* 83-98. Retrieved from http://www.psye.org/articulos/vista%20Santos%20Pastor.pdf
- Martínez, L. F., Castejón, F. J. & Santos, M. L. (2012). Diferentes percepciones sobre evaluación formativa entre profesorado y alumnado en formación inicialen educación física. Revista Electrónica Interuniversitaria de Formación del Profesorado, 15(4), 57-67. Retrieved from http://www.aufop.com/aufop/revistas/arta/digita1/171/1747
- Marton, F. & Säljö, R. (1976). On Qualitative Differences in Learning-2: Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology, 46*, 115-127. DOI: 10.1111/j.2044-8279.1976.tb02304.x
- Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in HigherEducation*, 31(2), 199-218. DOI: 10.1080/03075070600572090
- Palacios, A. & López-Pastor, V. M. (2013). Haz lo que yo digo pero no lo que yo hago: sistemas de evaluación del alumnado en la formación inicial del profesorado. *Revista de Educación*, *361*, 279-305. DOI: 10.4438/1988-592X-RE-2011-361-143
- Palacios, A., López-Pastor, V. M. & Barba, J. J. (2013). Tipologías de profesorado universitario en función de la evaluación aplicada a los futuros docentes. *Estudios sobre educación, 24*, 173-195. Retrieved from http://dspace.unav.es/dspace/bitstream/10171/2 9569/2/PALACIOS.pdf
- Pérez, A., Tabernero, B., López, V. M., Ureña, N., Ruiz, E., Capllonch, M. & Castejón, J. (2008). Evaluación formativa y compartida en la docencia universitaria: La concreción de cuestiones/clave para su aplicación en el camino hacia el Espacio Europeo de Educación Superior (EEES). *Revista de Educación*, 347, 435-451. Retrieved from http://www.mecd.gob.es/dctm/revista-de-educacion/articulosre347/re34720.pdf?documentdd=0901e72b8123677c

- Romero-Martín, R., Fraile-Aranda, A., López-Pastor, V. M. & Castejón-Oliva, F. J. (2014). Relación entre sistemas de evaluación formativa, rendimiento académico y carga de trabajo del profesor y del alumno en la docencia universitaria. *Infancia y Aprendizaje*, *37*(2), 310-341. DOI: 10.1080/02103702.2014.918818
- Rué, J. (2009). El aprendizaje autónomo en la Educación Superior. Madrid: Narcea.
- Sadler, E. R. (2005). Interpretations of criteriabased assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175-194. <u>DOI:</u> 10.1080/0260293042000264262
- Santos, M. L., Martínez, L. F. & López-Pastor, V. M. (Eds.) (2009). La innovación docente en el EEES. Experiencias de evaluación formativa y compartida en la formación inicial del profesorado. Universidad de Almería.
- Struyven, K., Dochy, F. & Janssens, S. (2005).

 Students' perceptions about evaluation and assessment in higher education: a review.

 Assessment & Evaluation in Higher Education, 30(4), 325-341.

 10.1080/02602930500099102
- Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of Further and Higher Education*, *33*, 57-69. DOI: 10.1080/03098770802638671
- Tejedor, F. J. (Coord.) (1998). Los alumnos de la Universidad de Salamanca. Universidad de Salamanca.
- Tonucci, F. (2010). Apuntes sobre el manifiesto. *Educar(nos), 49*, 9-11.
- Torrance, H. (2012). Formative assessment at the crossroads: conformative, deformative and transformative assessment. *Oxford Review of Education*, 38(3), 323-342. DOI: 10.1080/03054985.2012.689693
- Trillo, F. (2005). La evaluación de los estudiantes universitarios. El caso de la Universidad de Santiago de Compostela. Universidad de Santiago de Compostela.
- Vallés, C., Ureña, N. & Ruiz, E. (2011). La evaluación formativa en docencia universitaria. Resultados globales de 41 estudios de caso. *REDU*, 9(1), 135-158.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477-501. <u>DOI:</u> 10.1023/A:1023967026413

Romero-Martín, Rosario; Castejón-Oliva, Francisco-Javier & López-Pastor, Víctor (2015). Differences of students and faculty on the difficulties to implement the formative assessment. *RELIEVE*, 21 (1), art. ME5. DOI: 10.7203/relieve.21.1.5169

Acknowledgements

This paper is based on research developed under the R&D project: "*Teaching skills in the initial teacher education of physical education teachers*". Funding programme November 2013 of the National Programme for Research, Development and Innovation focused on Social Challenges under the framework of the State Programme for Scientific and Technical Research and Innovation 2013-2016. Reference: EDU 2013-42024-R.

Authors

To know more / Saber más

Romero-Martín, Rosario (rromero@unizar.es).

Professor at the University of Zaragoza. She is the author of contact for this article. Her research focuses on teaching physical activity, body language and emotions. His address is: Ronda Misericordia, 5. 22001 Huesca (Spain)







Castejón-Oliva, Francisco Javier (javier.castejon@uam.es).

His research lines are aimed at initial teacher of physical education, evaluation of learning and teaching the sport. He has several publications on student assessment in higher education. His address is: Universidad Autónoma de Madrid. Calle Francisco Tomás y Valiente, 3. Ciudad Universitaria de Cantoblanco. 28049 Madrid (Spain)





López-Pastor, Víctor (vlopez@mpc.uva.es).

Main lines of research: Formative and Shared Assessment in Higher Education, Training and Lifelong Initial Teacher, Formative Evaluation in Physical Education. His address is: Universidad de Valladolid. Facultad de Educación de Segovia. Campus María Zambrano (Universidad de Valladolid). Pza/ Alto Leones de Castilla, 1. 40005-Segovia (Spain)





Revista ELectrónica de Investigación y EV aluación Educativa E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]

- © Copyright, RELIEVE. Reproduction and distribution of this articles it is authorized if the content is no modified and their origin is indicated (RELIEVE Journal, volume, number and electronic address of the document).
- © Copyright, RELIEVE. Se autoriza la reproducción y distribución de este artículo siempre que no se modifique el contenido y se indique su origen (RELIEVE, volumen, número y dirección electrónica del documento).