

Comparison between conceptions and assessment practices among secondary education teachers: more differences than similarities (*Comparación entre concepciones y prácticas de evaluación en profesores de Educación Secundaria: más diferencias que semejanzas*)

Javier Fernández Ruiz¹ & Ernesto Panadero^{2,3}

Author Note

¹ Departamento de Psicología Evolutiva y de la Educación, Facultad de Psicología, Universidad Autónoma de Madrid, Spain.

² Facultad de Psicología y Educación. Universidad de Deusto, Spain.

³ IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Recommended citation:

Javier Fernández-Ruiz & Ernesto Panadero (2020) Comparison between conceptions and assessment practices among secondary education teachers: more differences than similarities (*Comparación entre concepciones y prácticas de evaluación en profesores de Educación Secundaria: más diferencias que semejanzas*), Journal for the Study of Education and Development, DOI: [10.1080/02103702.2020.1722414](https://doi.org/10.1080/02103702.2020.1722414)

This is a pre-print of an article published in Journal for the Study of Education and Development. Personal use is permitted, but it cannot be uploaded in an Open Source repository. The permission from the publisher must be obtained for any other commercial purpose. This article may not exactly replicate the published version due to editorial changes and/or formatting and corrections during the final stage of publication. Interested readers are advised to consult the official published version.

The final version can be accessed here:

<https://www.tandfonline.com/doi/full/10.1080/02103702.2020.1722414>

Corresponding author email: javier.fernandez.uam@gmail.com

Abstract: Teachers' conceptions can affect any teaching practice, including assessment. Compared with the previous learning stages in Spain, the accreditation-based focus increases in secondary education. Therefore, it is necessary to study the conceptions teachers have of assessment during this stage and the role these play in their teaching practices. The study participants comprised 219 secondary education teachers. A mixed methodology including a questionnaire, dilemmas and the analysis of test contents was used. An analysis was implemented of the relationship between the conceptions of assessment and the type of tests used, self-assessment and peer assessment, and characteristics of feedback. The teachers reported having a formative conception, but this was incoherent with their assessment methods characterised by a scarce use of self-assessment and peer assessment and a hardly formative feedback. In conclusion, we highlight the relevance of the teaching-related conceptions in the education system and the need for working jointly with teachers to implement an authentically formative assessment in the classroom, with an emphasis on consistency between conceptions and practices.

Keywords: Formative assessment, secondary education, teacher conceptions, feedback, self-assessment

The importance of teachers' conceptions as a determining factor in teaching processes has attracted increasing interest in recent decades (Monereo, Castelló, Durán & Gómez, 2009; Brown, 2009). Recent research studies have collected these conceptions (Brown, 2002; Remesal & Brown, 2015) and how they are constructed (Martín, Mateos, Martínez, Cervi, Pecharromán & Villalón; 2006). Furthermore, an exploration of their relationship with certain teaching practices has begun in recent years (Remesal, 2007, 2011; Brown, 2009). The purpose of this study is to analyse the conceptions of secondary education teachers in Spain with regards to assessment, and their relationship with the assessment and feedback methodologies teachers implement in their classrooms.

Teachers' conceptions about assessment

Goodenough (1990) defined conceptions as "...assertions on different aspects of reality that an individual assumes to be true at a given time in their life, without this meaning that they comprise an objective truth." Despite the lack of consensus on an exact definition, or about their degree of explicitness (Coll & Remesal, 2009), teachers' conceptions about teaching and learning comprise a prolific line of research that have helped us understand how to intervene in the classroom to improve education (Pajares, 1992). Conceptions change over the course of a teaching career and tend to be associated with specific teaching practices that encompass the entire teaching-learning process (Martín et. al., 2006), including assessment (Kahn, 2000). Given this relationship between teachers' conceptions and their teaching behaviours, it is essential to better understand what comprises the former to improve what occurs in the latter.

Teachers' conceptions about assessment are intimately tied with the functions that assessments assume in the education system: the improvement of teaching and learning, the certification of learning results and the accountability of teachers and schools (Brown, 2002). The first two goals, improvement and certification of learning, are problematic for teachers

because the means necessary to achieve these are frequently considered opposites. An initial split between conceptions emerges at this point: those inclined toward an assessment based on improving learning, and those inclined toward an assessment based on certifying the students' knowledge. These poles have been given different names in literature, with "formative assessment" as the most popular (also known as assessment for learning) vs. summative assessment (Black & Wiliam, 1998).

Furthermore, teachers' conceptions about assessment are associated with their conceptions about teaching and learning (Gordon, 2008). Those teachers with a traditional or deterministic vision of learning as their point of departure will consider assessment a summative tool. To the contrary, those teachers with a constructivist conception of learning will be inclined toward a formative conception of assessment (Eren, 2010).

Several authors have adapted this separation between Formative-Summative conceptions. Shortly after the publication by Black and Wiliam (1998), studies carried out in Spain already mentioned this polarisation (Coll & Onrubia, 1999). Outside of Spain, Brown (2002) was another author, among others, who used this bifurcation as the basis for a model of conceptions regarding assessment. Brown's model, despite having been supported in several later studies in different countries (Brown, 2007; Brown, Lake & Matters, 2009), is still pending confirmation with a Spanish sample (Remesal, 2009). Specifically, this study uses one of the few models that have been tested with a sample in Spain (Remesal & Brown, 2015). Its complete classification is as follows.

<< Insert Table 1 here >>

As already mentioned, teachers' conceptions may be associated with their classroom practices (e.g., Brown, 2009), particularly with regards to assessment (Quesada-Serra, Rodríguez-Gómez & Ibarra-Sáiz, 2016). While a summative conception of assessment will translate into a methodology aimed at assessing learning through objective measurements (for

example, assessment based strictly on tests, scarce feedback from the teacher, etc.), a formative conception will entail a methodology that promotes learning (for example, self-assessment or peer assessment activities, more detailed and informative feedback, etc.) (Black & Wiliam, 2003).

Formative assessment

Black and Wiliam (2009) define formative assessment as follows:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (p. 9)

This definition lays the theoretical foundations upon which formative assessment is built. However, the term has been a subject of discussion in recent years. Authors like Gordon Stobart (2008) point out a frequent misinterpretation of the term when a series of tests are applied over the course of the instructional process. However, the fact that assessment is ongoing does not necessarily make it formative. For assessment to be formative, it must provide sufficient information upon which to base decision-making aimed at improving a student's learning.

Once the concept and its implications are defined, indicators are needed that will enable a teacher to understand which specific aspects must be addressed in practice. Wiliam and Thompson (2007) describe five key strategies linked to formative assessment:

1. Clarifying and sharing learning intentions and criteria for success.
2. Engineering effective classroom discussions, questions, and learning tasks.
3. Providing feedback that moves learners forward.
4. Activating students as instructional resources for one another.
5. Activating students as owners of their own learning.

When these are implemented in secondary education, they become methodologies that encompass the entire assessment process, from designing and presenting activities, to using thorough, detailed feedback. With the goal of considering the last four key points of Wiliam and Thompson (2007), this study focuses on three specific elements: the design and structure of the evidence for assessment, the characteristics of the feedback offered, and the use of self-assessment and peer assessment. These are explored in-depth below.

Design and structure of the evidence for assessment.

Often, the evidence used the most in the classroom have not been designed with a teaching-related purpose, wherefore their use on an individual level does not usually offer students information that enables them to improve (Black & Wiliam, 2003). One example of this are final exams as the only evidence; these provide information on the result, but not the process. In fact, a national study carried out in Spain shows that exams are still the most commonly used evidence in higher education (Panadero, Fraile, Fernández Ruiz, Castilla-Estévez & Ruiz, 2018), possibly due to the influence of academic regulations for universities. Because of this, to expand formative information for students, evidence of assessment has become more advanced (for example, learning portfolios o diaries) and is combined and used throughout the entire process instead of only at the end.

However, we must understand that the use of some evidence of assessment in particular does not necessarily imply that the assessment is formative. This may actually be appreciated in the design of the assessments, given that specific evidence (a test, for example) may be designed to assess only memory-based learning or capture significant learning processes.

Alonso-Tapia and Villa-Arocena (1996) analysed the characteristics of evidence of assessment used in secondary education in Spain. Their study demonstrated that, depending on the field of study, the evidence had unique and highly differentiated characteristics.

Twenty years ago, these authors found that the predominant assessment in social or humanities-related subjects (language, literature, social sciences) was based on recall and the comprehension of facts and concepts by using short questions. To the contrary, in science-based studies (mathematics, physics, chemistry), assessment was based on applying principles for problem-solving. This division is contrary to the goals of formative assessment, in which both concepts and procedures and attitudes should be assessed across all areas of the school curriculum (Alonso-Tapia & Villa-Arocena, 1996).

The study by Alonso-Tapia and Villa-Arocena (1996) reveals historic weaknesses in the design of evidence. However, formative assessments encompass many other aspects; if our goal is for a student to learn from the assessment, it is also necessary for the feedback provided on the student's learning to be as thorough and personalised as possible.

Characteristics of feedback.

Feedback, defined as information offered by an agent (teacher, peer, or oneself) on a student's performance (Hattie & Timperley, 2007) is an indispensable tool for improving learning. However, perhaps because of the process's complexity, feedback is not always provided adequately. Figure 1 displays a general overview:

<< Insert Figure 1 here >>

One of the objectives of formative assessment is to reduce the discrepancies between real performance and desired goals. There are two basic ways in which this objective may be achieved: to modify the goals so that they are more appropriate for the student, and to offer students strategies that will enable them to attain said goals. Information on a student's performance is what both processes require, precisely. In this regard, formative feedback seeks answers to three main questions (Hattie & Timperley, 2007): "Where is the student?" (*Feed Back*), "Where is the student headed?" (*Feed Up*), and "How can the student get

there?" (*Feed Forward*). Each one entails a specific type of feedback, and only by combining these can students be offered exhaustive information that will enable their improvement.

Yet, the feedback offered in secondary education tends to be far from the model proposed by Hattie and Timperley. This feedback usually focuses on answering the first question and rarely refers to adapting or achieving goals. Authors such as Bernard (2000; in Amaranti, 2010) point out the scarce clarity of the feedback that is commonly used, which tends to comprise a numeric score or qualitative terms, such as "insufficient", "sufficient" or "outstanding", which are hardly informative for students.

Furthermore, as reflected in the definition of Black and Wiliam (2009), the sole responsibility for providing feedback is not assumed by the teacher, but rather the students and their classmates themselves should assume an active role in the evaluation process. Wiliam and Thompson (2007) also point out the importance of activating students as instructional resources for both themselves and their classmates. This implies empowering the students and capacitating them for autonomous learning, in addition to being able to learn from their classmates and allowing them, in turn, to learn. Self-assessment and peer assessment are essential tools for this.

Use of self-assessment and peer assessment.

Self-assessment is one of the key strategies of formative assessment, given its use for promoting students' self-regulation (Nicol & McFarlane-Dick, 2006; Panadero & Alonso-Tapia, 2013) and for facilitating learning (Brown & Harris, 2013). However, despite the empirical evidence that demonstrates how the use of self-assessment favours self-regulated learning (Panadero, Jonsson & Botella, 2017), the strategy is still hardly used in the classroom (e.g., Panadero, Brown & Courtney, 2014). This is also true for peer assessment. Using peers as evaluators has proven to be a key strategy in formative assessment (Sadler,

2009). However, it is scarcely used in secondary education, and in those cases when it has been applied, it is usually superficially and without continuity (Panadero & Brown, 2017).

To date, we are familiar with many formative assessment practices, and efforts have been made recently to implement it in the classroom. Nevertheless, all change processes represent major challenges for teachers (Monereo, Weise & Álvarez, 2013). Therefore, despite our familiarity with the theory, it is necessary to be knowledgeable of the regional contexts to propose adequate changes, given the contextual variations, and in terms of conceptions that are hardly generalisable (Black & Wiliam, 2005). Few studies in our country have samples recruited from secondary education.

Spain's historical context

As a result of the instability of legislation on education in recent decades, the Spanish context is especially interesting (Panadero & Brown, 2017). In this regard, we must point out the major change that occurred in our education system in 1990. Organic Law 1/1990, dated 3 October, on the LOGSE (*Ley Orgánica de Ordenación General del Sistema Educativo*, General Plan for the Education System) came into force that year, and entailed major structural changes that greatly impacted the secondary education teachers that were teaching then.

Among other things, the LOGSE established a vision of assessment as a formative tool, something never proposed before in Spain. In addition to decreasing the teachers' opportunities for direct decision-making on students' passing courses, it also established the cycle-based system, which enables teachers to implement activities adapted to students with difficulties when passing from one academic year to another. We must recall that, given the duration of a teacher's professional career, approximately half of the current secondary education teachers began teaching before the LOGSE became effective. The magnitude of the changes that this law generated, and the ones that followed, prompt us to foresee differences

in the teachers' conceptions, depending on their experience and the period in which they began teaching, as previous studies have shown (Castejón & Martínez, 2001).

Research questions

Based on the aforementioned issues, the purpose of our study was to answer three questions (H = hypothesis):

1. What conceptions about assessment do secondary education teachers have, and how are these associated with their teaching experience?
 - H1: Years of teaching experience are positively correlated with the conception of summative assessment.
2. What is the relationship between these conceptions and the assessment methodology used in the classroom?
 - H2: The conception of formative assessment is associated with: (H2a) a greater variety of evidence of assessment reported by the teachers, (H2b) a greater variety of tests used (understood as diversity of contents, processes and types of questions included in a test) and (H2c) a greater use of self-assessment and peer assessment.
3. What is the relationship between the conceptions and the characteristics of feedback reported by the teachers?
 - H3: The conception of formative assessment is correlated with a greater use of *Feed Up* and *Feed Forward* in the teacher's responses.

Method

Participants

The sample comprises 219 secondary education teachers from 15 autonomous communities, with a mean experience of 18.68 years ($SD = 10.162$). Of these, 75 were selected based on the criterion of proximity with the researchers and the remaining participants were chosen randomly. Table 2 displays the general composition of the sample.

<< Insert Table 2 here >>

Procedure

Before data collection began, a pilot study was implemented with four teachers who answered the questionnaire and dilemmas in the researcher's presence. The participants and researcher jointly analysed in depth the responses to the feedback-related dilemmas. Finally, to ensure the instruments' viability, concerns and impressions about the entire data collection process were collected for their subsequent analysis.

Data collection took place in two blocks. First, four centres with which the researcher had connections were contacted. Teachers from these four centres (a total of 75) completed the paper version of the questionnaire. The remaining participants were recruited after creating a database of all the schools in provincial capitals in Spain and randomly contacting their directors, who distributed an on-line version of the questionnaire among their teachers.

Data collection instruments

Teachers' conceptions.

The teachers' conceptions of assessment were measured using the questionnaire by Remesal and Brown (2015, p. 339). This questionnaire comprises 34 items using Likert-type scales and evaluates two referent conceptions regarding the purpose of the assessment (formative regulation and societal control), and four referents regarding the focus of assessment (teaching, learning, certifying and accountability). Although the original questionnaire design included an answer scale with positive scoring and 6 response options (4 positive, 2 negative), our study used a Likert-type scale from 1-5. This design of including a neutral option was chosen to avoid potential response bias by the teachers (Abad, Ponsoda, Olea & García, 2011).

Furthermore, the teachers were requested to position themselves along a continuum, with 0 representing a completely summative conception and 100 representing a completely

formative conception, and after that to position the conceptions they believed were predominant in each stage of education (primary, secondary and university).

Evidence of assessment.

Two tools were used to study the evidence of assessment. The first was a questionnaire that measures the use of five different types of evidence: tests, tasks, homework, and portfolios, among others, as well as their characteristics, and about the person in charge of scoring them, to detect the use of self-assessment and peer assessment (more details are given in Appendix A).

The second tool contains examples from real tests used by the participants. All of the teachers included in the sample were requested to send a sample test by e-mail to the authors, though only 24 (11% of the total) did so. Most of the tests analysed belong to the field of Humanities (12) or Science (9), in addition to two tests from the Social Sciences and another from Art. The analysis of these tests recorded their characteristics as regards contents, procedures and formal structure, according to the categories established by Alonso-Tapia and Villa-Arocena (1996). The tests were coded dichotomously, depending on the presence or absence of each category (see Appendix C for further details).

Dilemmas on feedback.

Two dilemmas designed ad hoc were used, one related with *Feed Up* and another with *Feed Forward* (see Appendix B for further details). These dilemmas requested open-ended responses of the teachers which were coded dichotomously, depending on the presence or absence of *Feed Up* or *Feed Forward* in their feedback. Inter-rater reliability was assessed to verify the reliability of the response categories, obtaining Kappa scores of .74 ($p < .001$) in the first dilemma and .83 ($p < .001$) in the second.

Data analysis

To compare the hypotheses, SPSS software was used and the Student *t* test performed on hypotheses 2c and 3. Furthermore, an ANOVA was performed on hypothesis 3, the Pearson correlation coefficient on hypotheses 2a and 2b, and the linear regression analysis on hypothesis 1. With regards to the qualitative data obtained from the feedback dilemmas, a thematic analysis was done of the teachers' responses to identify unique practices. For their subsequent quantitative analysis, inter-rater reliability was assessed and the data were coded in numerical and dichotomous categories (presence/absence of *Feed Up/Feed Forward*).

Results

What conceptions about assessment do secondary education teachers have, and how are these associated with their teaching experience?

As Table 3 shows, the teachers were more inclined toward a formative conception of assessment than a summative conception. As regards assessment, the teachers are inclined toward assessment focused on learning ($M = 3.475$), over those focused on teaching ($M = 3.142$), accountability ($M = 3.152$) or certification ($M = 3.157$).

<< **Insert Table 3 here**>>

The teachers were asked directly to rate their conceptions on a scale from 0 to 100, with 0 representing a completely summative conception and 100 representing a completely formative conception. Furthermore, they were asked to use the same method to score the conceptions they believed predominated in each stage of education. The participants obtained a mean score of 60.34 on the formative-summative continuum, as is shown in Table 4. To the contrary, the mean score for secondary education was 51.67. In other words, they considered themselves to be a mean of 8.67 points above their colleagues in the same stage of education as regards their formative conception of assessment. In addition, 73.9% of the teachers in secondary education consider their own conceptions to be equally or more formative than

those of their colleagues. Finally, the teachers seem to detect a gradual decrease of the formative nature of the conceptions with an increase of the education stage.

<< Insert Table 4 here >>

Regarding the impact of years of experience on the teachers' conceptions of assessment, the linear regression analysis reveals a significant, though slight, relationship, between years of teaching experience and summative conception ($r^2 = .028$; $p = .024$). This result supports hypothesis 1.

What is the relationship between the conceptions and the assessment methodology used in the classroom?

A great majority of the teachers declared using tests, tasks and homework for purposes of assessment in their classrooms. As regards tests, most teachers declare using 2 or 3 (29.2% and 21.1%, respectively) per trimester, implying between 6 and 9 tests per academic year. In addition, most of the teachers report using both individual and group tasks, the former being more popular (86.8%, compared with 67.6%). Likewise, many teachers claim to use different methods for evaluating homework, whether reviewing the students' notebooks (69.4%) or asking them to perform a task on the blackboard (73.1%).

Less popular than the former types of evidence are the learning portfolio or diary, which only 34.2% of the teachers report using. Finally, 21.5% of the participants declare to use evidence of assessment different from the aforementioned, including oral exams, project-based learning or laboratory practices. None of these types of evidence had a presence exceeding 5%, wherefore they were omitted in the subsequent analyses. We must mention, however, that because we did not present the teachers a description of each type of evidence, it is possible that some may have declared using a given evidence of assessment which in reality was not the case, as a result of misinterpreting the terms.

As Table 5 displays, though there are differences in the amounts of evidence used according to the teachers' conceptions, these were not significant, therefore failing to support hypothesis 2.a which claimed that the formative conception would be associated with a greater variety of evidence of assessment according to teachers' reports. However, worth mentioning are the existing differences in those who used homework as an assessment tool, in that the use of homework had a significant correlation with the questionnaire's "Learning" scale.

<< Insert Table 5 here >>

The tests that were remitted by 24 participants were analysed in terms of amount of contents, processes and types of questions they included. Though there were not any significant differences as regards the variety of the contents, there were significant differences as regards the amount of psychological processes assessed. As Figure 2 shows, the psychological processes evaluated and the types of questions used show a subtle increase in teachers with scores above the mean in formative conception, while the opposite occurs in teachers with a summative conception. Specifically, differences in relation to the summative subscale become significant after performing the Mann-Whitney test ($p = .007$ and $p = .043$, respectively). These results partially support hypothesis 2.b, which claimed that a formative conception of assessment would be related with a greater variability in the tests used, understood as diversity of contents, processes and types of questions included in a test. Given the small sample size, however, these results must be interpreted with caution.

<< Insert Figure 2 here >>

As Table 6 shows, the use of self-assessment and peer assessment differs, depending on the evidence of assessment used. While tests are almost completely evaluated by the teacher, tasks are more frequently evaluated using peer assessment (particularly for group

tasks) and self-assessment. Self-assessment is also frequently used: 17.8% used it to evaluate homework, and 25.3% used it to evaluate portfolios.

<< Insert Table 6 here >>

As Table 7 shows, teachers' conceptions impact the teachers' reported use of self-assessment and peer assessment, with the latter being used significantly less by teachers with a summative conception of assessment, in support of hypothesis 2.c.

If we analyse the evidence individually, we can verify that the self-assessment in tests and tasks is used significantly less by teachers with a summative conception of assessment. If we focus on the use of peer assessment in tasks, it increases significantly in teachers with a formative conception of assessment, and decreases significantly in teachers with a summative conception of assessment.

Worth highlighting is how the use of self-assessment and peer assessment decreases notably in those teachers with a high score in the "Teaching" scale, likely due to the wariness of integrating this type of assessment with the official curricula. Perhaps for the same reason, teachers with a high score in the "Accountability" scale use self-assessment significantly less.

<< Insert Table 7 here >>

What is the relationship between the conceptions and the characteristics of feedback reported by the teachers?

Regarding the characteristics of feedback, the responses to the dilemmas reflected a relatively common use of *Feed Forward* (67.2%). To the contrary, the responses that included the use of *Feed Up* are less common, used by 72 teachers (37.9%).

After our review of the answers the teachers provided, we must point out the scarce formative level, in general, of the feedback they report. This is especially visible in the responses given to the dilemma linked with *Feed Up*, the majority of which are characterised

for their succinctness and scarce information for the student. The following cases that include the teacher's complete answer are an example of feedback of poor quality:

"I would tell the student that they must learn to define the space dedicated to each section on their own. This is an inappropriate error at the high school level." F. One year of experience.

"I would tell the student that they had not adhered to the given specifications and did not follow the instructions given in class." P. Two years of experience.

"I would explain to the student that it's great that they wished to further explore the author's life, but that they should have focused on what was requested by the task." V. Seven years of experience.

"It is important to differentiate between the details and the main issue." G. Twenty-nine years of experience.

In general, the responses given to the second dilemma are more constructive.

However, a significant portion of responses were left incomplete, which would be hardly useful for a student in a classroom setting.

"Learning consists of making a note of the errors committed and correcting them on subsequent occasions in which similar situations are proposed." G. Twenty-nine years of experience.

"I would ask the student if they understood what I had said, and I would repeat my explanation and ask the student again if they had understood." G. Fifteen years of experience.

Regarding the relationship between feedback type and teaching conceptions, as shown in Table 8, the use of *Feed Up*, and especially *Feed Forward*, in the responses to the dilemmas increases in those teachers with a formative conception of assessment. However, the significance of the former case is not enough to enable generalisation to the population. These results do not reject hypothesis 3.

Altogether, the results obtained after combining the results of the two dilemmas show that the combined use of *Feed Up* and *Feed Forward* in the dilemmas increases significantly in those teachers with a formative conception of assessment. As regards the focus of the assessment, we must point out the results concerning the “Learning” scale, given that the teachers included in the sample with a high score in this scale used both *Feed Up* and *Feed Forward* more frequently. Nevertheless, the results cannot be generalised because the required significance level was not reached.

<< Insert Table 8 here >>

Discussion

The three main objectives of this study were: to analyse the conceptions of secondary education teachers regarding assessment, to analyse the relationship between these conceptions and the assessment methodology they used in the classroom, and to explore the relationship between their conceptions and the feedback they offered their students. The conclusions drawn from the results obtained are presented below.

Concerning the first objective, the secondary education teachers reported having a greater formative, rather than summative, conception of assessment. Given the differences in the response scale completed by our teacher sample, it is difficult to compare these results with those obtained by Remesal and Brown (2015) with teachers of *español como lengua extranjera* (ELE, Spanish as a foreign language). However, we observe that secondary education teachers, as is the case with ELE teachers, many of whom work in non-formal education settings, show similar results regarding conceptions of assessment, favouring formative over summative assessment.

The responses teachers give when asked about the conceptions they believe predominate in a given stage of education are also striking. We must mention that these questions about “conceptions about conceptions”, given the absence of a common definition

of the term “conception”, cannot assure the validity of the responses. However, the teachers’ responses are coherent with previous literature. Authors such as Panadero, Brown and Courtney (2014) already reveal a decrease of the formative focus of conceptions in higher education levels, in favour of a greater certifying nature. The fact that secondary education teachers are aware of this phenomenon, which according to their responses seems to be the case, could impact the way in which they receive their students who have completed primary education and how they prepare them for access to the university.

Furthermore, we find that those teachers with more years of teaching tend to report more summative conceptions, a result also found in previous studies (Castejón & Martínez, 2001). One hypothesis is that these differences in the teachers’ conceptions could have been impacted by the influence of changes in legislation. If this is true, the practice of seasoned teachers who began teaching before the LOGSE was implemented could be impacted by the legislative framework and professional training that were previously in effect. Nevertheless, the relationship between conceptions and teaching experience could be affected by numerous factors in addition to legislation on education.

Regarding the second objective, to explore the relationships between conceptions and practices of assessment, we must highlight the apparent disconnection between conceptions and practices reported by the teachers, given that their high mean scores on the questionnaire's formative scale led us to expect that they would implement formative practices. However, the results show the opposite, given that the assessment methodologies the teachers reported and the tests they use reveal that most of them use a hardly formative approach, as was already observed in other studies with Spanish samples (Alonso-Tapia & Villa-Arocena, 1996; Panadero & Brown, 2017).

In particular, the 24 tests analysed show scarce variety of types of contents and psychological processes assessed, as many of their questions were limited to assessing the

lowest cognitive processing levels of the taxonomy proposed by Bloom (e.g., recall and comprehension). Furthermore, the majority of the tests followed the same format: long responses and text commentaries for Humanities-related subjects and computation and problem-solving for Science-related subjects. These results, despite having been extracted from a small sample, could demonstrate a stagnation in the design of evidence of assessment, given that Alonso-Tapia and Villa-Arocena (1996) found very similar results more than two decades ago.

In addition, the teachers of our sample also report a scarce use of self-assessment and peer assessment. Though their use varies, depending on the evidence, in general, 34.8% of the teachers use self-assessment, and 30.0% use peer assessment. These rates are even lower than those obtained in previous studies with Spanish secondary education teachers, in which higher percentages of teachers reported using self-assessment (Panadero et al., 2014) and peer assessment (Panadero & Brown, 2017). These differences could be explained by the format of the question posed in this study in referring to a habitual use, while the majority of the teachers in the aforementioned studies affirmed using self-assessment and peer assessment only occasionally. We must mention that neither is the students' participation fostered in assessing processes in higher stages of education (Rodríguez-Gómez, Ibarra-Sáiz & García-Jiménez, 2013).

Regarding our third and final objective, similar results are found regarding the relationship between the conceptions and the feedback given. The teachers, despite reporting more formative conceptions, reflect a feedback of more summative characteristics in their responses to the dilemmas. Though the use of *Feed Forward* in these responses is relatively frequent, the same is not the case for *Feed Up*, which is just as necessary to offer quality feedback (Hattie & Timperley, 2007). Furthermore, the analysis of the teachers' responses reveals that an important proportion of feedback is omitted, superficial or hardly formative.

These results are similar to those obtained internationally, in which the feedback most commonly used in the classroom focused on the most basic levels (task and personal), compared with those that are more productive for learning (process and self-regulation) (Hattie, 2009). Students also perceive this low formative quality of feedback, especially in the last courses of secondary education (Núñez et al., 2015; Katz, Kaplan & Gueta, 2009).

Clarifications about the disconnection between conceptions and practices

Until now, we have referred to the disconnection between conceptions and practices in general, given that the teacher's reported having formative conceptions, but their practices of assessment did not present this focus. The assessment methodologies they claim to use, the tests analysed and the feedback collected in the dilemmas are clearly summative. However, beyond the descriptive level, and as was the case in previous studies (Brown, 2009), we found differences in the teachers' assessments, depending on their conceptions.

Teachers with a summative conception seem to design tests that entail fewer psychological processes and types of questions. In addition, these teachers tend to report a use of less complete feedback and lower frequency of methodologies that would enable the students' empowerment. This combination allows us to maintain our hypothesis that the teachers with a summative conception of assessment use a methodology that could be considered less pedagogical, according to the criteria of Wiliam and Thompson (2007).

Given that seasoned teachers tend to have conceptions that are more summative, and that these are reflected in an assessment methodology that is even less formative than that of their more novel colleagues, we must ask ourselves which activities are necessary to improve the formative quality of assessment in secondary education. This study proposes two areas of intervention for training teachers and designing standardised tests.

Implication for teacher training

The first would be to pay greater attention to ongoing teacher training, which currently seems to have little effect on participating teachers (Gómez & Guerra, 2012). Furthermore, this training should focus on the most relevant aspects of formative assessment, considering the learning assessed, the evaluator and the quality of feedback.

Ongoing training of teachers is a resource that has already been used successfully in the attempt to improve teachers' assessment practices, as previous experiences have shown (Panadero & Brown, 2017). However, offering teachers training on formative assessment strategies, though necessary, would be insufficient unless the focus of standardised tests is also changed, which is the second area of intervention we propose.

Implications for the design of standardised tests

Ultimately, teachers implement teaching and assessment strategies in coherence with the educational system where they teach, and the context of secondary education with tests of a markedly certifying nature, like the EvAU, represents an important disadvantage in the attempt to implement formative assessment in the classrooms (Laborda, 2012). In this sense, we must draw attention to the apparent contradiction of our government's discourse on education. Though recent legislation has promoted formative assessment and competency-based learning, tests in schools continue to have a strong summative nature.

Important progress is being made in this field, an example of which are the diagnostic tests in the Basque Country (Basque Government, 2015), with a methodology that shows major efforts to evaluate student progress so as to foster their learning. Yet, these types of tests are still an exception to the rule, and the first step along a long road yet to be travelled.

Limitations

Of the three most relevant limitations, the first is the fact that data collection was based on self-reports. Given the absence of direct observation of assessment practices, the participants' responses may have been affected by the social desirability bias. With the goal

of addressing this limitation, we attempted to analyse real tests that would complement the data collected through self-reports. However, these tests represent a small part of the sample.

Second, the lack of consensus among researchers as to the definition of the term “conceptions” and the degree to which they are organised and hierarchical could lead to other researchers’ considering that we have not evaluated conceptions, as these can only be measured in practice. Concerning our perspective on this issue, which is shared by many researchers, using a questionnaire allows for accessing a more cognitive part of the teachers’ conceptions. Furthermore, we presented dilemmas as the data collection method, which allows us to reflect on the more implicit aspect of teachers’ conceptions, to a certain extent.

Third, the generalisation of the results may be affected because we did not control for the sample’s representativeness as we did not select public-private-private schools with state-funded financial support nor schools from different autonomous communities.

Conclusion

Our results show contradictions because although the teachers declare having formative conceptions of assessment, they report practices that are more summative. Furthermore, those teachers with more summative conceptions report using assessment methodologies with a more summative approach, here in greater coherence with their conceptions. Therefore, the assessments of learning that are taking place in secondary education seem to be distanced from the theoretical and empirical tenets of formative assessment. Though with limitations, this data reveals the need for implementing efforts jointly with teachers, with the goal of providing them the theoretical and practical tools they need to implement real formative assessment in their classrooms.

References

- Alonso-Tapia, J; Villa-Arocena, J. L. (1996). ¿Cómo se Evalúa el Aprendizaje en Enseñanzas Medias? *Revista de Ciencias de La Educación*, 168, 473–503.
- Amaranti, M. (2010). Concepciones y prácticas de retroalimentación de los profesores de lenguaje y comunicación de primer año de educación media: Investigación cualitativa con estudio de caso. *Congreso Iberoamericano de Educación*, 0–22.
- Black, P., y Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: Department of Education and Professional Studies, Kings College.
- Black, P., y Wiliam, D. (2003). “In Praise of Educational Research”: Formative assessment. *British Educational Research Journal*, 29(5), 623–637. JOUR.
<http://doi.org/10.1080/0141192032000133721>
- Black, P. y Wiliam, D. (2005) Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices, *The Curriculum Journal*, 16(2), 249–261.
- Black, P., y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. JOUR. <http://doi.org/10.1007/s11092-008-9068-5>
- Brown, G. T. L. (2002) *Teachers' conceptions of assessment*. (Tesis doctoral). University of Auckland, Auckland, NZ
- Brown, G. T. L. (2007, December). Teachers' conceptions of assessment: Comparing measurement models for primary & secondary teachers in New Zealand. In *annual conference of the New Zealand Association for Research in Education, Christchurch, NZ*.
- Brown, G. (2009). Teachers' self-reported assessment practices and conceptions: Using structural equation modelling to examine measurement and structural models. *Structural equation modelling in educational research: Concepts and applications*.
- Brown, G. T., Lake, R., & Matters, G. (2009). Assessment policy and practice effects on New Zealand and Queensland teachers' conceptions of teaching. *Journal of Education for Teaching*, 35(1), 61-75.

- Brown, G. T. L., y Harris, L. R. (2013). Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks, CA: SAGE
- Castejón, J.L. y Martínez, M.A. (2001). The personal constructs of expert and novice teachers concerning the teacher function in the Spanish educational reform. *Learning and Instruction, 11*, 113-131.
- Coll, C., & Remesal, A. (2009). Concepciones del profesorado de matemáticas acerca de las funciones de la evaluación del aprendizaje en la educación obligatoria. *Infancia y aprendizaje, 32*(3), 391-404.
- Eren, A. (2010). Consonance and dissonance between Turkish prospective teachers' values and practices: Conceptions about teaching, learning, and assessment. *Australian Journal of Teacher Education, 35*(3), 3.
- Gobierno Vasco. (2015). *Evaluación Diagnóstica*. Obtenido de Evaluación diagnóstica: http://ediagnostikoak.net/edweb/cas/materiales-informativos/ed15cas/ED15_modelo_cas.pdf
- Gómez, V., y Guerra, P. (2012) Teorías implícitas respecto a la enseñanza y el aprendizaje: ¿Existen diferencias entre profesores en ejercicio y estudiantes de pedagogía? *Estudios Pedagógicos, 38*(1), 25-43.
- Goodenough, W. H. (1990). Evolution of the Human Capacity for Beliefs. *American Anthropologist, 93*, 597-612.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie, J., y Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.
- Katz, I., Kaplan, A., & Gueta, G. (2009). Students' needs, teachers' support, and motivation for doing homework: A cross-sectional study. *The Journal of Experimental Education, 78*(2), 246-267.
- Laborda, J. G. (2012). Presentación. De la Selectividad a la Prueba de Acceso a la Universidad: pasado, presente y un futuro no muy lejano. *Revista de Educación (357)*, 17-28.

Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo. Boletín Oficial del Estado, num. 238, de 4 de Octubre de 1990, pp. 28927 a 28942.

<https://www.boe.es/boe/dias/1990/10/04/pdfs/A28927-28942.pdf>

Martín, E., Mateos, M., Martínez, P., Cervi, J., Pecharromán, A. y Villalón, R. (2006). Las concepciones de los profesores de educación primaria sobre la enseñanza y el aprendizaje. En J. I. Pozo, N. Scheuer, M.P. Pérez Echeverría, M. Mateos, E. Martín y M. de la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje*, (pp. 171-187). Barcelona: Graó.

Monereo, C., Castelló, M., Duran, D., & Gómez, I. (2009). Las bases psicoeducativas del proyecto PISA como guía para el cambio en las concepciones y prácticas del profesorado de secundaria. *Infancia y aprendizaje*, 32(3), 421-447.

Monereo, C., Weise, C., y Álvarez, I. (2013). Cambiar la identidad docente en la universidad. Formación basada en incidentes dramatizados. *Infancia y Aprendizaje*, 36(3), 323-340.

Nicol, D. J., y Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. JOUR. <http://doi.org/10.1080/03075070600572090>

Núñez, J. C., Suárez, N., Rosário, P., Vallejo, G., Cerezo, R., & Valle, A. (2015). Teachers' feedback on homework, homework-related behaviors, and academic achievement. *The Journal of Educational research*, 108(3), 204-216.

Onrubia, J., & Salvador, C. C. (1999). Evaluación de los aprendizajes y atención a la diversidad. In *Psicología de la instrucción: la enseñanza y el aprendizaje en la educación secundaria* (pp. 141-168). Horsori.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of educational research*, 62(3), 307-332.

Panadero, E. y Alonso-Tapia, J. (2013). Revisión sobre autoevaluación educativa: Evidencia empírica de su implementación a través de la autocalificación sin criterios de evaluación, rúbricas y guiones. *Revista de Investigación en Educación*, 11(2), 172-197.

- Panadero, E., Brown, G., y Courtney, M. (2014). Teachers' reasons for using self-assessment: a survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy & Practice*, 21(4), 365–383. article. <http://doi.org/10.1080/0969594X.2014.919247>
- Panadero, E., y Brown, G. T. L. (2017). Teachers' reasons for using peer assessment: positive experience predicts use. *European Journal of Psychology of Education*, 1–24. article. <http://doi.org/10.1007/s10212-015-0282-5>
- Panadero, E., Jonsson, A., y Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98.
- Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2018). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, 1-19.
- Quesada-Serra, V., Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2016). What are we missing? Spanish lecturers' perceptions of their assessment practices. *Innovations in Education and Teaching International*, 53(1), 48-59. doi:10.1080/14703297.2014.930353
- Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: The Spanish instance, building upon Black and Wiliam. *The Curriculum Journal*, 18(1), 27-38.
- Remesal, A. (2009). Spanish student teachers' conceptions of assessment when starting their career. In *Symposium: Perceptions and conceptions of assessment in the classroom: Different national perspectives. 13th Conference of the European Association for Research in Learning and Instruction*.
- Remesal, A., y Brown, G. T. L. (2015). Conceptions of assessment when the teaching context and learner population matter: compulsory school versus non-compulsory adult education contexts. *European Journal of Psychology of Education*, 30(3), 331–347.
- Rodríguez-Gómez, G., Ibarra-Sáiz, M. S., & García-Jiménez, E. (2013). Autoevaluación, evaluación entre iguales y coevaluación: conceptualización y práctica en las universidades españolas. *Revista de Investigación en Educación*, 11(2), 198–210.

Sadler, D. R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In *Assessment, learning and judgement in higher education* (pp. 1-19). Springer, Dordrecht.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.

William, D., y Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? En C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Erlbaum.

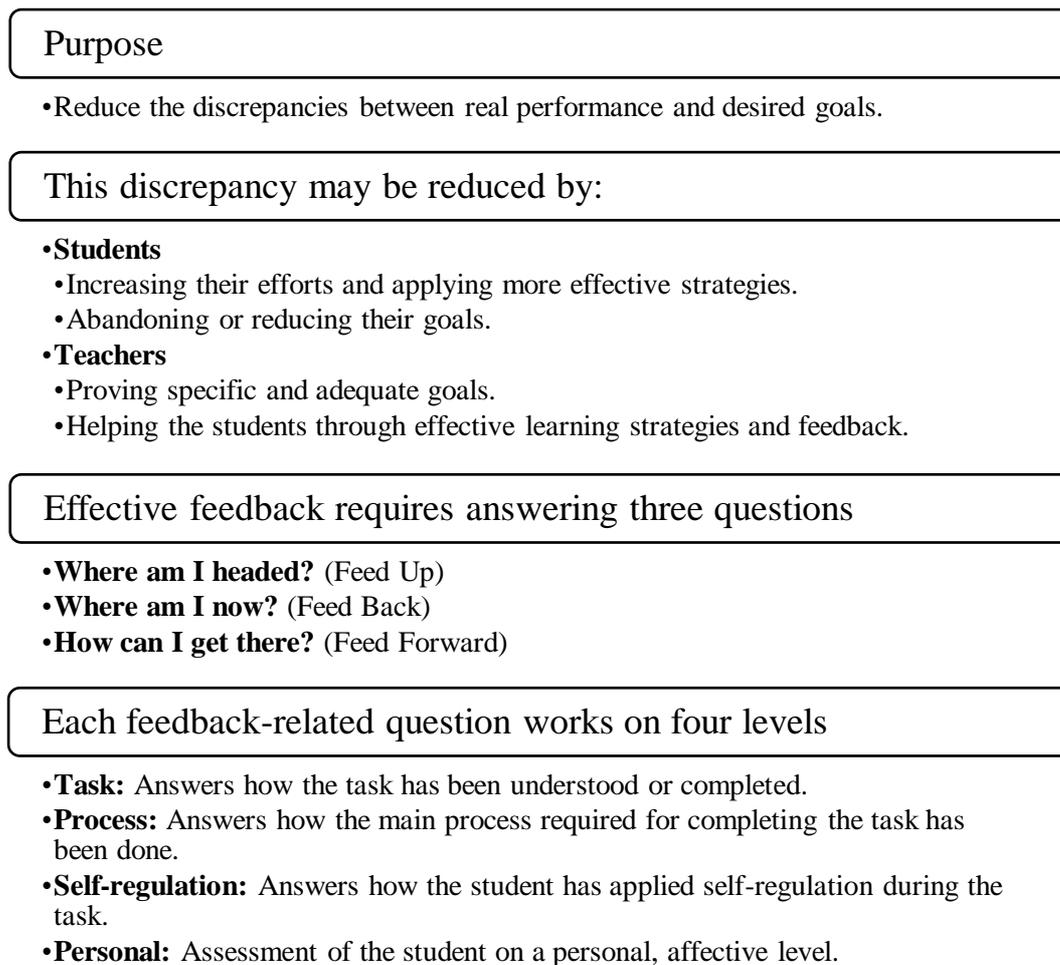


Figure 1. Feedback model for fostering learning (Hattie & Timperley, 2007, p. 87).

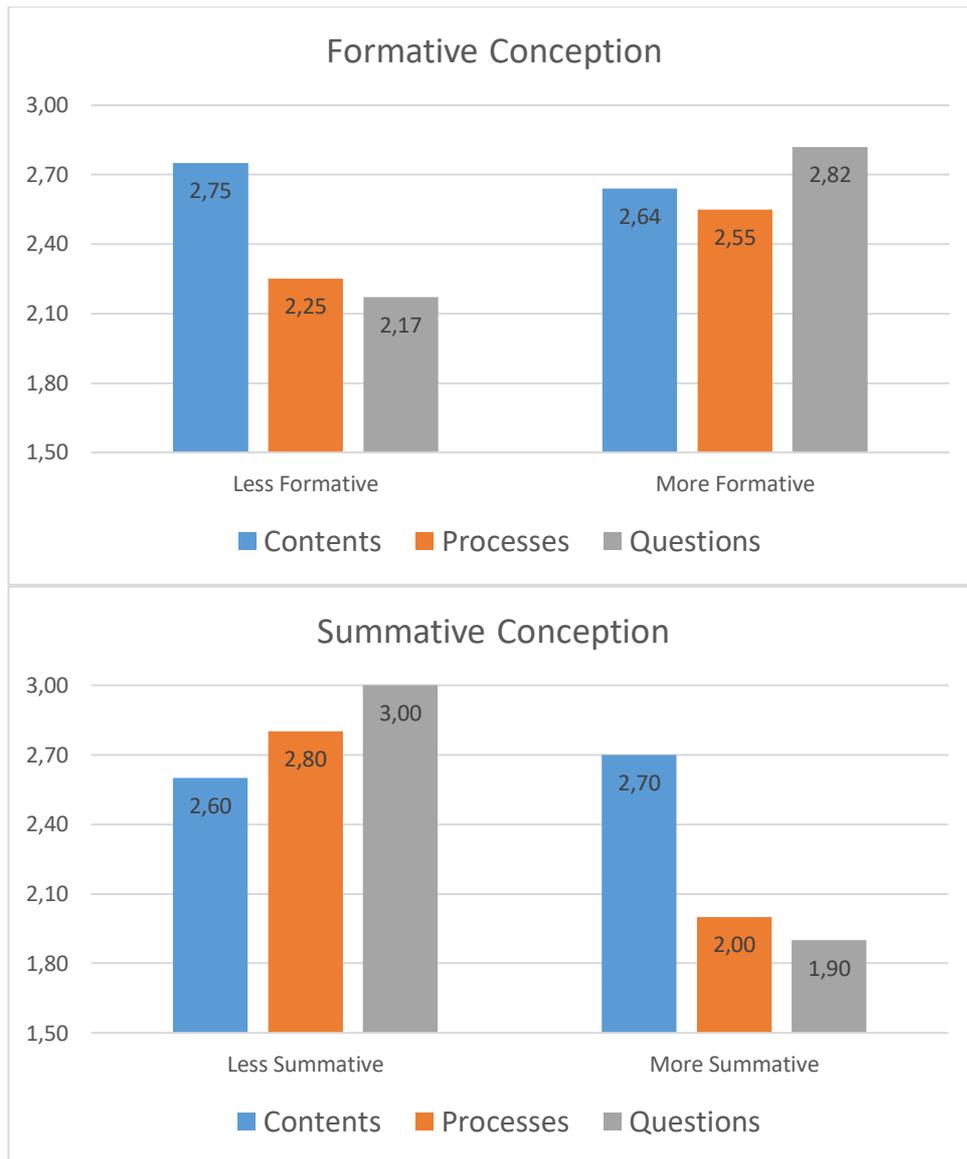


Figure 2. Type of tests with data grouped according to teachers' conceptions

TABLE 1

Classification of the teachers' conceptions of assessment (Remesal & Brown, 2015)

Regarding the function of assessment in the system

| | |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| Formative regulation: | Assessment should be used with the goal of promoting and improving the students' learning. This may be considered a formative conception of assessment. |
| Societal control: | Assessment should be used with the goal of measuring and certifying the students' learning. This may be considered a summative conception of assessment. |

Regarding the assessment's focus of attention:

| | |
|------------------------|----------------------------------------------------------------------------------------------------------|
| Learning: | Related with the effects of assessment on the students in terms of their motivation and self-regulation. |
| Teaching: | Related with the integration of assessment with the teaching curriculum and strategies. |
| Accountability: | Related with the use of assessment for testing and grading students. |
| Certifying: | Related with the use of assessment for establishing education standards, scholarships and awards. |

TABLE 2
Distribution of the sample

| | | <i>M</i> | <i>SD</i> |
|---------------------------------------------------|--------------------------|----------|-----------|
| Years of teaching experience | | 18.68 | 10.162 |
| | | <i>K</i> | % |
| Field of teaching | Sciences | 76 | 34.9 |
| | Humanities | 67 | 30.7 |
| | Social sciences | 29 | 13.3 |
| | Art | 12 | 5.5 |
| | Physical education | 4 | 1.8 |
| | Other | 30 | 13.8 |
| Teaches at the Baccalaureate level ⁽¹⁾ | Yes | 122 | 56 |
| | No | 96 | 44 |
| Experience in a non-formal setting ⁽²⁾ | Yes | 136 | 62.4 |
| | No | 82 | 37.6 |
| Autonomous community | Principality of Asturias | 66 | 30.1 |
| | Aragon | 25 | 11.4 |
| | Castille and Leon | 24 | 10.9 |
| | Community of Madrid | 21 | 9.6 |
| | Andalusia | 6 | 2.7 |
| | Extremadura | 4 | 1.8 |
| | Catalonia | 2 | 0.9 |
| | Murcia | 2 | 0.9 |
| | Navarra | 2 | 0.9 |
| | Cantabria | 1 | 0.5 |
| | Castille and La Mancha | 1 | 0.5 |
| | Community of Valencia | 1 | 0.5 |
| | La Rioja | 1 | 0.5 |
| | Melilla | 1 | 0.5 |
| | No response given | 62 | 28.3 |

(1) Teaches at the Baccalaureate level every year or almost every year. In addition, may teach at the obligatory secondary level (ESO).

(2) Has taught in a non-formal setting (academies, private lessons, etc.) at some point of the teaching career.

TABLE 3
Conceptions of assessment

| <i>Purpose</i> | <i>M</i> | <i>SD</i> |
|----------------------------|----------|-----------|
| Formative regulation | 3.800 | 0.383 |
| Societal control | 2.718 | 0.404 |
| <i>Focus of assessment</i> | | |
| Teaching | 3.142 | 0.299 |
| Learning | 3.475 | 0.324 |
| Certifying | 3.157 | 0.315 |
| Accountability | 3.152 | 0.460 |

TABLE 4
Conceptions of assessment reported directly by the teachers

| | <i>M</i> ⁽¹⁾ | <i>SD</i> |
|----------------------|-------------------------|-----------|
| Personal conceptions | 60.34 | 20.007 |
| Primary education | 65.98 | 24.346 |
| Secondary education | 51.67 | 19.458 |
| University education | 40.36 | 26.228 |

| <i>Personal conceptions vs. Conceptions in secondary education</i> | | |
|--------------------------------------------------------------------|----------|----------|
| | <i>N</i> | <i>%</i> |
| + Formative | 82 | 37.4 |
| = to Formative | 80 | 36.5 |
| - Formative | 36 | 16.4 |

⁽¹⁾100 = Completely formative conceptions; 0 = Completely summative conceptions

TABLE 5
Relationships between Conceptions and Evidence of assessment

| <i>Purpose</i> | Amounts of evidence | | Tests | | Tasks | | Homework | | Portfolio | |
|----------------------------|---------------------|----------|----------|----------|----------|----------|--------------|-------------|-----------|----------|
| | Corr. (1) | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> |
| Formative regulation | .115 | .114 | -1.035 | .302 | 0,633 | .528 | 0,988 | .345 | 1,240 | .216 |
| Societal control | -.105 | .148 | 0,831 | .429 | -1,072 | .285 | 0,269 | .789 | -1,653 | .100 |
| <i>Focus of assessment</i> | | | | | | | | | | |
| Teaching | -.057 | .432 | -0,172 | .863 | -1,255 | .211 | 0,272 | .790 | -0,716 | .475 |
| Learning | .032 | .654 | 0,371 | .711 | -0,807 | .420 | 2,988 | .003 | -0,195 | .846 |
| Certifying | -.034 | .635 | -0,013 | .989 | 0,385 | .701 | 0,174 | .862 | -0,019 | .984 |
| Accountability | .030 | .671 | -0,494 | .622 | 0,020 | .984 | 0,662 | .509 | -0,405 | .686 |

Note: Statistically significant results are displayed in bold

(1) Corr. = Pearson correlation coefficient

TABLE 6
Distribution of evidence of assessment

| | K | % | Number of tests | | Evaluator (%) | | |
|-----------|-----|------|-----------------|------------|---------------|-------------------|-------------------|
| | | | M | SD | Teacher | SA ⁽¹⁾ | PA ⁽²⁾ |
| Tests | 208 | 95 | 2.75 | 1,146 | 100 | 12.4 | 6.2 |
| | | | Type (%) | | | | |
| | | | Individual | Group | | | |
| Tasks | 209 | 95.4 | 86.8 | 67.6 | 99.5 | 15.3 | 21.1 |
| | | | Notebooks | Blackboard | | | |
| Homework | 205 | 94 | 69.4 | 73.1 | 94.6 | 17.8 | 7.4 |
| Portfolio | 75 | 34.2 | | | 89.3 | 25.3 | 9.3 |
| Other | 47 | 21.5 | | | 87.2 | 13.0 | 19.6 |

1. SA = Self-assessment
 2. PA = Peer assessment

TABLE 7
Use of self-assessment and peer assessment

| | Total | | | | Tests | | | | Tasks | | | |
|----------------------|-------------------|-------------|-------------------|-------------|---------------|-------------|----------|----------|---------------|-------------|---------------|-------------|
| | SA ⁽¹⁾ | | PA ⁽²⁾ | | SE | | PA | | SE | | PA | |
| | Corr. | <i>p</i> | Corr. | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> |
| Formative regulation | .048 | .514 | .133 | .068 | 0,267 | .789 | 0,401 | .689 | 1,235 | .219 | 3,200 | .002 |
| Societal control | -.192 | .008 | -.299 | .000 | -2,446 | .015 | -1,276 | .204 | -3,662 | .000 | -4,915 | .000 |
| Teaching | -.135 | .061 | -.139 | .055 | -1,606 | .110 | -0,379 | .705 | -1,526 | .129 | -1,698 | .091 |
| Learning | -.083 | .246 | -.099 | .164 | -1,754 | .081 | -1,017 | .310 | -1,483 | .140 | -0,639 | .523 |
| Certifying | -.057 | .425 | -.113 | .116 | -0,448 | .655 | -0,709 | .479 | -0,912 | .363 | -1,136 | .257 |
| Accountability | -.154 | .030 | -.136 | .056 | -2,149 | .033 | -1,357 | .177 | -2,533 | .012 | -0,918 | .360 |

Note: Statistically significant results are displayed in bold

(1) SA = Self-assessment

(2) PA = Peer assessment

TABLE 8
Relationship between conceptions, Feed Up and Feed Forward

| | <i>Feed Up</i> | | <i>Feed Forward</i> | | Total | |
|----------------------------|----------------|----------|---------------------|-------------|--------------|-------------|
| | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> |
| <i>Purpose</i> | | | | | | |
| Formative regulation | 1.940 | .054 | 2.646 | .009 | 5.267 | .006 |
| Societal control | -0.698 | .486 | -0.407 | .684 | 0.286 | .751 |
| <i>Focus of assessment</i> | | | | | | |
| Teaching | 1.308 | .193 | 0,678 | .499 | 0,900 | .408 |
| Learning | 1,933 | .055 | 1.672 | .096 | 2.888 | .058 |
| Certifying | 0.003 | .998 | 0.514 | .608 | 0.147 | .863 |
| Accountability | -0.043 | .966 | 1.360 | .176 | 0.962 | .384 |

Note: Statistically significant results are displayed in bold

Appendix A. Evidence of assessment

1. Please specify which of the following evidence of assessment you use habitually with your classes during a single trimester:

- a. **Tests**..... ()
 - i. Number of tests (approx.)..... ()
 - ii. Scorer
 - 1. Teacher..... ()
 - 2. Classmates..... ()
 - 3. Self-assessment..... ()
- b. **Tasks**..... ()
 - i. Task type
 - 1. Individual..... ()
 - 2. Group..... ()
 - 3. Both..... ()
 - ii. Scorer
 - 1. Teacher..... ()
 - 2. Classmates..... ()
 - 3. Self-assessment..... ()
- c. **Homework**..... ()
 - i. How it is reviewed
 - 1. Review of students' notebooks..... ()
 - 2. Task performed at the blackboard/during class..... ()
 - 3. Both..... ()
 - ii. Scorer
 - 1. Teacher..... ()
 - 2. Classmates..... ()
 - 3. Self-assessment..... ()
- d. **Learning portfolio/diary**..... ()
 - i. Scorer
 - 1. Teacher..... ()
 - 2. Classmates..... ()
 - 3. Self-assessment..... ()
- e. **Other**..... ()
 - i. Please specify which: _____
 - ii. Scorer
 - 1. Teacher..... ()
 - 2. Classmates..... ()
 - 3. Self-assessment..... ()

Appendix B. Dilemmas

Dilemma 1. – FEED UP

This dilemma is designed on the basis of defining *Feed Up* as an answer to the question, “Where am I headed?” (Hattie & Timperley, 2007). A case is presented of a student that is given instructions to complete a task, but has not clearly understood the desired goals before starting, wherefore the final result is imbalanced. Therefore, the feedback given should be related to goal-setting in subsequent tasks.

The students are asked to complete a task, individually, about an important author. It is explained to them that the tasks must include a brief biography of the author and an analysis of the author’s most important contribution. A clear limit of 5 pages is set.

When reviewing that tasks, you discover that one of the students has written a very long biography of the author, with details and reflections that far exceed the instructions. As a result, the student did not have sufficient space to analyse the author's main contribution, and only addressed this superficially in the last two paragraphs of the task.

In your own words, explain what feedback you would provide this student.

Dilemma 2. – FEED FORWARD

This dilemma is designed upon the understanding that *Feed Forward* answers the question, “How can the student get there?” (Hattie & Timperley, 2007). It poses a case in which a student that had been given feedback on another task previously should know later what they had done incorrectly. Given that this student commits the same error in a subsequent task, it is expected for the feedback given to refer to specific instructions about how to resolve these errors, what we understand as *Feed Forward*.

In the same task as in the previous case, you discover that another student had drafted the task in a confusing way, in which the main ideas are unclear. When sharing this comment with the student, the student agrees with the assessment and promises to do better next time.

However, when submitting the next task (similar in nature to the previous one), you discover that the student has the same problems as the first time. The task is again confusing to read, and the ideas are still unclear.

In your own words, explain what feedback you would provide this student.

Appendix C. Coding of tests (Alonso-Tapia & Villa-Arocena, 1996)

TYPES OF CONTENT the student is expected to know:

- Facts: occurrences, events, processes (like a chain of facts), etc.
- Concepts: notions, definitions, categories, characteristics comprising a concept, etc.
- Procedures: steps to be implemented, mechanisms of action for...
- Principles: rules, models, theories, formulas, etc.

TYPES OF PSYCHOLOGICAL PROCESSES that the student must perform to correctly answer what is being asked:

- Recall: searching for and recalling information stored in one's memory and recognising the information presented.
- Comprehension: building a functional representation of the information: categorising, paraphrasing, summarising, abbreviating, etc.
- Prediction: given a situation, anticipating or extrapolating what will occur.
- Application: using the acquired knowledge to: - analyse and summarise new information, - manipulate, - decide on the course of action chosen to...
- Evaluation: comparing situations, objects or events with a series of criteria given for decision-making.

TYPES OF QUESTIONS or formal structure of these:

- Essay: open-ended questions to be addressed, thematic.
- Short open-ended questions: the student must provide a brief answer that includes very specific contents.
- Text commentary: analyse a text's parts and structure, functions, ideas and/or conclusions.
- Objective tests: the student must choose an answer from among several options.
- Problem-solving: the student must interpret the content expressed in a text in which a problem is posed and reorganise and/or code that information to perform the operations necessary to find a solution.
- Computation: this is differentiated from the preceding category in that the starting point is already coded in mathematical language and the student merely needs to complete a mechanical operation, according to the formulation of the question (simplification, reduction, operation, etc.).