

**Assessment design decisions in practice: Profile identification in approaches  
to assessment design.**

Javier Fernández Ruiz<sup>1</sup>, Ernesto Panadero<sup>2,3</sup>, Daniel García Pérez<sup>4</sup> & Leire Pinedo<sup>2</sup>

<sup>1</sup> Departamento de Psicología Evolutiva y de la Educación, Facultad de Psicología, Universidad Autónoma de Madrid, Spain.

<sup>2</sup> Facultad de Psicología y Educación. Universidad de Deusto, Spain.

<sup>3</sup> IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

<sup>4</sup> Departamento de Psicología. Universidad Europea de Madrid

**Corresponding author:**

Javier Fernández Ruiz

Aula PDIF, Facultad de Psicología

Universidad Autónoma de Madrid

Email: javier@fernandezruiz.com

**Research funded by:** Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad) National I+D Call (Convocatoria Excelencia) project reference EDU2016-79714-P; and personal grant (Formación de Personal Investigador) reference BES-2017-080054.

**Recommended citation:**

Fernández Ruiz, J., Panadero, E., García-Pérez, D. & Pinedo, L. (2021) Assessment design decisions in practice: Profile identification in approaches to assessment design, *Assessment & Evaluation in Higher Education*, DOI: [10.1080/02602938.2021.1937512](https://doi.org/10.1080/02602938.2021.1937512)

This is a post-peer-review, pre-copyedit version of an article published in *Assessment & Evaluation in Higher Education*. The final authenticated version is available online at: <https://doi.org/10.1080/02602938.2021.1937512>. This manuscript 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 published version.

The authors declare to not having any conflict of interest regarding this manuscript. This research has been approved by the ethics committee from Universidad Autónoma de Madrid.

### **Abstract**

This study aims to identify different profiles of higher education teachers based on the way they design their assessment methods. It also explores differential characteristics in each profile and differences in the assessment methods preferred by the teachers and those implemented in their subjects. Sixteen teachers from four universities participated carrying out a think-aloud simulation task: designing the assessment methods for a set of learning outcomes. Teachers' testimony during the task was transcribed and categorized using content analysis and an open-coding procedure. Three different patterns were identified: (a) focus on the feasibility of the assessment tasks, (b) on the alignment with the learning outcomes, or (c) alignment with teaching methods. Most of the participants focused only in one of the three elements. Teachers also designed different assessment methods in the simulation task in comparison with the ones they use in their subjects, despite the guidelines of assessing the same learning outcomes. A lack of resources is claimed as the reason of these differences. Implications for future research are discussed.

**Keywords:** *Assessment design; higher education; instructional design; assessment*

## **Introduction**

Much has been studied on how different assessment practices can promote beneficial effects on students (Boud & Falchikov, 2006; Andrade & Heritage, 2013; McMillan & Moore, 2020), but educational practices are not usually consistent with empirical knowledge about assessment (Ibarra-Sáiz & Rodríguez-Gómez, 2010). In higher education, teachers can make decisions about several aspects of their assessment practices, and yet not much is known about the teacher's perspective when designing and implementing those practices (Bearman et al. 2017). The studies that address this issue have mostly focused on self-reported data (Bennett et al. 2011; Norton et al. 2013; Bearman et al. 2017), and literature exploring the assessment design while it takes place is missing. A better understanding of assessment design processes would be beneficial to increase the consistency between the wide empirical knowledge about assessment and the practices taking place in the classroom. Using an innovative data collection and analysis methodology, this study aimed to shed more light on the way in which university teachers design their assessment methods.

## **Theoretical framework**

### **Assessment design in higher education**

There are many studies discussing how specific assessment practices, if done properly, can have beneficial effects for students on aspects such as self-regulation (Panadero & Romero, 2014), motivation (Cauley & McMillan, 2010), and learning (Black & Wiliam, 1998). Although there is agreement on the importance of assessment as a promoter of learning, its implementation in higher education is sometimes far from what we know as formative assessment (Wu & Jessop, 2017). Studies such as Panadero et al. (2019) and Jessop & Tomas (2017) show a traditional assessment panorama in European universities. This is not necessarily due to a lack of effort on the part of

university teachers, who spend a third of their time designing and promoting assessment activities (Izci & Caliskan, 2017). However, it has been argued that, due to an inadequate understanding and use of assessment techniques, higher education teachers do not end up obtaining benefits in relation to teaching and learning (Stiggins, 2007).

To fully understand this phenomenon, it is necessary to investigate the way that teachers design their assessment practices, to identify challenging areas that may be stopping them from achieving a set of assessment practices they feel comfortable with (Bearman et al. 2017). The empirical understanding of this process would be able to lead to professional actions aiming to increase assessment effectiveness and make it more beneficial for teachers and students.

In this vein, authors such as Carless (2015) consider assessment design as one of the most critical aspects of assessment practices. It is known that this process can be extremely complex for teachers, balancing multiple tensions at the personal and institutional level. McDonald & Joughin (2009) and Meyer et al. (2010) have produced conceptual works representing the factors influencing assessment design. These factors go from general government and institutional policies to student and teacher characteristics. Several authors agree that assessment design by teachers is characterised by a tension between contextual and personal factors (McMillan, 2003; Bearman et al. 2017).

On one hand, contextual influences may be related to the subject, the discipline (Meyer et al. 2010), the learning outcomes, and the requirements of the department. The organizational culture is also frequently mentioned because teachers from the same faculty or department end up adapting their assessment practices to fit those of their peers (Carless, 2015; Bearman et al. 2017).

On the other hand, personal factors such as past experiences (Panadero et al., 2014), backwash effects (Watkins et al. 2005), teaching experience (Quesada-Serra et al. 2016), and conceptions about assessment (Fernández Ruiz & Panadero, 2020) regulate assessment design. Brookhart (2011) recommends a series of skills for teachers in relation to the design of assessment practices, which include being able to articulate clear learning objectives, communicate to students what achievement of a learning objective looks like, understand the purposes and uses of the range of available assessment options and be skilled in using them, and having the skills to analyse classroom questions, test items and performance assessment tasks.

Research, however, shows a different picture in higher education assessment practices. The assessments implemented are often traditional and out of step with current trends (Jessop & Tomas, 2017; Panadero et al., 2019), and teachers do not seem to be able to justify the assessment practices they have designed (Postareff et al. 2012). Quesada-Serra et al. (2016) show how teachers, despite showing their agreement regarding the importance of formative assessment practices, do not use them in their classrooms with the same frequency (Norton et al. 2005).

These discrepancies are an indicator of teachers' difficulties when it comes to achieving coherence between their aims and their practices regarding assessment. It is therefore conceivable that, at some point in the design, teachers encounter certain challenges that they are not capable of facing, and that prevent them from carrying out assessment practices as they would like.

### **Assessment design processes**

Although the factors affecting assessment design have been a relatively common research topic during recent years, much less is known about the design process itself.

Bearman et al. (2017) categorised assessment design actions into three different types: essential, selective, and meta-design.

The essential design activities are those actions which must be undertaken in order to design and implement an assessment task. These actions can be, for example, the development of an assessment task, feedback processes, and alignment with the teaching methods. Selective activities have to do with the grading processes, the control over contract cheating, or the fairness of the assessment practices. Finally, meta-design activities refer to the activities that educators undertook to manage their own design processes across iterations. Examples of this type would be the periodic review of assessment practices, collaboration with other teachers, or the active search for feedback from students.

In summary, it is known that educational assessment is a major key to ensure student achievement in higher education. However, the assessment practices used at this stage tend to be different from what the teachers would like to implement, which makes it necessary to investigate more carefully how the assessment design process is carried out. Assessment design is an extremely complex action, and the variety of different factors involved might make it difficult to recall in a further interview. Therefore, it is important to explore this process while it is taking place. This study aimed to investigate the way in which university teachers design their assessment practices using data collection methods new to this topic based on the observation and analysis of a simulated assessment design. To do so, we start from three research questions:

RQ1 – Are there different teacher assessment design profiles?

RQ2 – Are there differential characteristics in the teachers belonging to each of the profiles?

RQ3 – Which are the differences between the assessment practices designed during the assignment and the assessment practices implemented in their subjects?

## Method

### Participants

The participants were 17 higher education teachers, belonging to the branches of sport sciences, medicine, and mathematics. Two additional teachers were discarded from the analysis as they offered incomplete results. The data regarding academic discipline, years of experience, and gender is listed in the Table 1.

**Table 1.**

*Information about the participants.*

	Gender	University	Faculty	Teaching experience (years)	Current unit experience (years)
1	Male	1	Sport Sciences	31	16
2	Male	1	Sport Sciences	20	8
3	Male	1	Mathematics	20	3
4	Female	4	Sport Sciences	6	4
5	Male	2	Mathematics	31	1
6	Teaching pair (male + female)	1	Medicine	15	3
7	Male	1	Mathematics	28	2
8	Male	3	Sport Sciences	17	13
9	Female	2	Medicine	15	5
10	Male	1	Mathematics	28	1
11	Female	4	Sport Sciences	10	7
12	Male	2	Medicine	38	25
13	Female	1	Mathematics	6	2
14	Male	2	Mathematics	20	6
15	Male	4	Sport Sciences	14	11
16	Male	4	Sport Sciences	10	7

### Procedure

The present study used a qualitative approach to explore the assessment process in a practice-based setting. It sought to describe and explain the phenomenon within a

professional context. To cover the range of data needed in this study and to obtain an intensive focus of the data, the sample size was limited.

Participants were contacted by email after analysing their syllabus as part of a previous study now submitted for publication (see Fernández Ruiz et al., submitted for publication), which intended to explore and compare the normal assessment patterns in each faculty. After this analysis, the research team contacted the participants, having the focus on a balanced sample in terms of teaching experience and academic discipline. An individual session was held which included an initial interview and a simulation task. Before starting with the data collection, a pilot study was performed with three participants not included in the final sample. The goal was to ensure that the interview questions and the simulation task were understandable and to anticipate possible questions by the final participants.

During the initial interview, participants were asked about contextual factors such as teaching experience and training, and about the development of their assessment practices over the years. The questions asked were the following:

1. Why did you choose these assessment practices instead of any others?
2. Did you carry out modifications in the assessment practices of this subject?  
Which were the reasons for those modifications?
3. How did your approach to assessment changed with the pass of the years?

Once the interview was completed, the participants were informed about the characteristics of the simulation task. The aim of the task was to show as faithfully as possible how teachers design their assessment practices. Each teacher was given a set of contents and learning outcomes, taken from the syllabus of their actual subjects to ensure that they were familiar to them. Then, the researcher asked them to design an assessment methodology to assess 'in the best way possible' these learning outcomes.

Participants were informed that there were no correct or incorrect ways to perform the task, and that replicating the assessment methods they use in their subjects was a valid option. A thinking aloud protocol procedure was followed, and the participants were instructed to make explicit all the ideas and thoughts that they had during the process. The task was generally well-received among the participants, as the pilot study helped to anticipate their possible concerns about it (e.g., Do I have to use different assessment methods in this task? What if I would consider different approaches depending on the classroom size or characteristics?). The reflections of the teachers were audio recorded and transcribed for later analysis. The teachers took between 15 and 30 minutes to complete the task.

### **Data analysis**

All the interviews were transcribed and coded using a content analysis approach. For RQ1, an open coding process was followed for the different actions that teachers performed during the task. After the analysis of the presence and order of the actions, three different profiles were found and described. Using the profiles as closed categories, two researchers independently read the transcriptions and coded the teachers into one of the profiles. This process had an initial agreement percentage of 90%, and subsequently the cases with discrepancies were discussed until they reached an agreement. Descriptive statistics were performed for RQ2, including frequencies, means, and standard deviations, and an open coding process was used for the interview responses. For RQ3, a closed coding process was followed using the categorization by Bearman et al. (2017). Software tools Atlas.ti 8 and SPSS 26 were used during the data analysis stage.

## **Results**

### **RQ1 – Are there different teacher assessment design profiles?**

Actions performed by the teachers during the task were coded into eight categories, described in Table 2.

**Table 2.**

*Assessment design process categorization.*

<b>Evidence</b>	Teachers selected one assessment evidence.
<b>Grading</b>	Teachers reflected on how the evidence must be graded, its weight on the students' final grade, and the minimum score required for passing.
<b>Learning outcomes</b>	Teachers read and analysed the learning outcomes provided and explored possible links with their assessment practices.
<b>Teaching methods</b>	Teachers selected one or several teaching methods and explored possible links with their assessment practices.
<b>Feedback</b>	Teachers reflected on how feedback must be provided to students.
<b>Rubrics</b>	Teachers explore the use of rubrics as a tool for students or themselves.
<b>Contract cheating</b>	Teachers explore the possible ways of preventing contract cheating in their assessment practices.
<b>Students' participation</b>	Teachers explore the possible ways of implementing students' participation in their assessment practices, mostly via self- or peer assessment.

Every action may have been carried out by the teacher once or more times during the task, and each time it would be coded independently. The presence or absence of the actions and the order in which they are carried out were used to find possible profiles among the participants. After the content analysis of the teachers' actions, three different profiles were seen in the design of the assessment practices. We called these profiles Classic, Competence, and Cohesive.

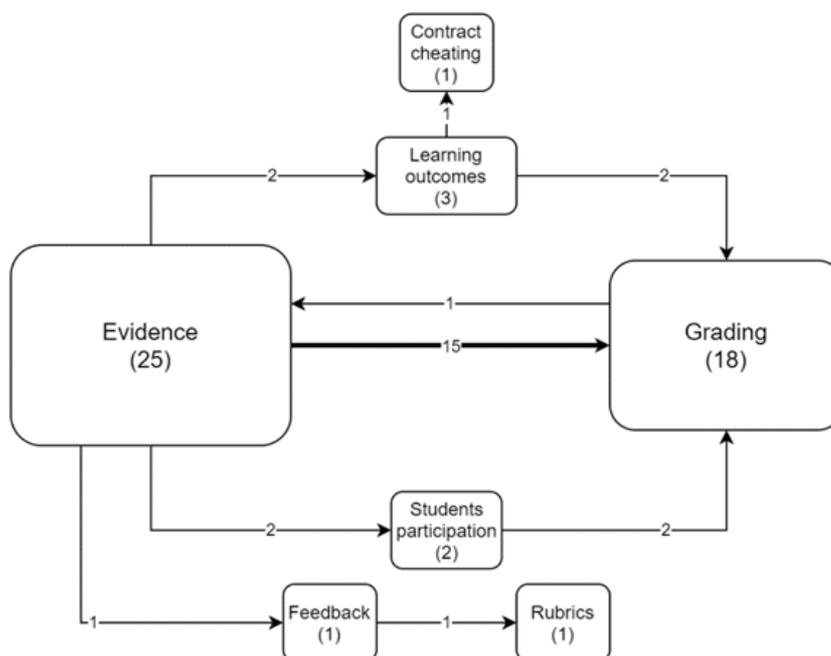
### **Classic profile**

A total of 7 teachers were categorised in the classic profile (Figure 1), designing a total of 25 instruments. In this profile, teachers designed their assessment practices in the most minimalist way possible. The teachers of this profile prioritised the efficiency and logistics of the assessment practices above all else. During the task, they tended to think directly about the assessment instruments they considered most convenient,

without placing too much emphasis on the competences assessed by each one of them. For each instrument, they usually indicated aspects such as the grading processes (common in all profiles), the temporality, the security regarding contract cheating, and the available resources.

**Figure 1.**

*Classic profile flowchart.*



**Note:** Numbers inside the boxes represent how many occasions the category was used. Numbers in the arrows represent how many occasions the design process went from one category to the next one.

As shown in Figure 1, 25 assessment instruments were designed in this way. Most of the teachers went on to give them a percentage of the grade, without further consideration. On two occasions they questioned the best way of preventing contract cheating, and on two other occasions whether they could involve the students in some way.

*It depends on how you want to name it, but it would be reasonable that the partial examinations have a grade. When that grade exceeds the threshold, we*

*are going to make, then the subject is passed. What would I do? According to my experience these years... to make the weight of each partial examination progressive. The first one will be 25% and so on until we have 100%.*

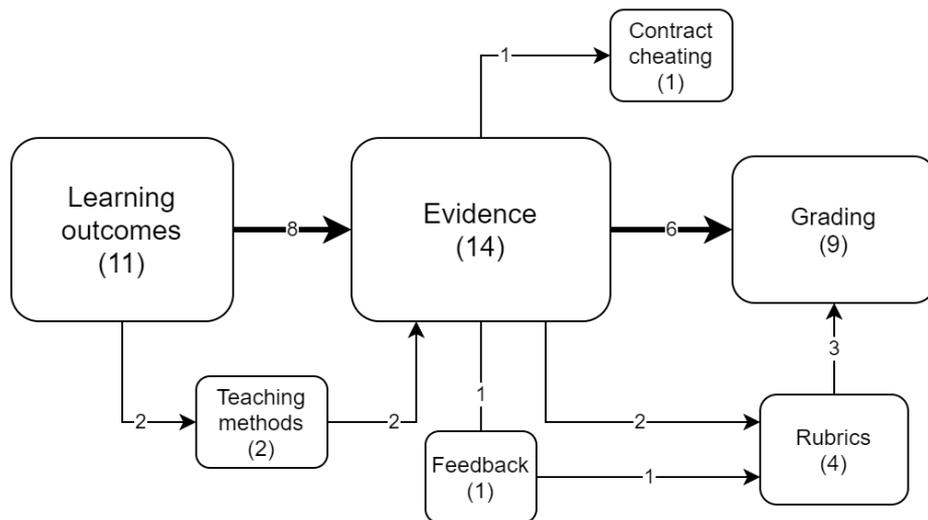
*Teacher 3 - Mathematics*

### **Competence profile**

A total of 5 teachers were categorised as competence profile, designing a total of 15 assessment instruments as shown in Figure 2. The teachers of this profile seemed to give priority to the coherence between the contents and learning outcomes to be assessed, and the chosen assessment instruments.

**Figure 2.**

*Competence profile flowchart.*



**Note:** Numbers inside the boxes represent how many occasions the category was used. Numbers in the arrows represent how many occasions the design process went from one category to the next one.

The most common pattern of this type of teacher was to take a close look at the learning outcomes of the subject, reflect on the best way to elicit them in an assessment task (either the one they currently use or some other proposal), and describe the assessment task they would use and how they would grade it. However, it is relatively

common for them to make a pause to consider the feedback provided or the use of rubrics.

*If we want the students to be able to collect and interpret relevant data to make judgements [one learning outcome from her subject] ... the assessment tasks would be the ones I am using right now, an assignment and a final examination. That makes them learn how to search information, or even using the information they already have in their notes, and make a judgement, a reflection, to answer one particular question.*

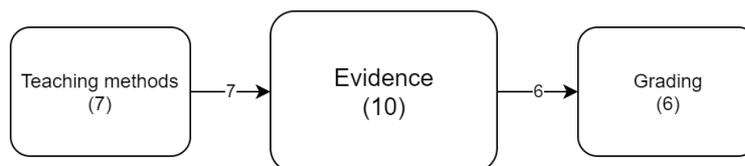
*Teacher 11 - Sport Sciences*

### **Cohesive profile**

A total of 4 teachers were categorised as cohesive, designing a total of 10 assessment instruments. Teachers of this profile seemed to pay special attention to integrating their assessment practices in a broader context. Normally, they tended to think first of the teaching methods that they would implement in their subject, after which the chosen assessment practices were a logical consequence. Teachers in this category are also characterised by being especially attentive to coordination between subjects, or even between different academic courses.

### **Figure 3.**

*Cohesive profile flowchart.*



**Note:** Numbers inside the boxes represent how many occasions the category was used. Numbers in the arrows represent how many occasions the design process went from one category to the next one.

*My teaching methodology is based on case-solving and problem-based learning (..). I would link to that one skills seminar, where the student can show their skills. It does not have to be only clinical exploration, but calculus, critical reading, and writing as well. All of it would be incorporated into the assessment element.*

*Teacher 12 - Medicine*

The syllabi of each teacher's subjects were analysed to explore possible differences in the assessment practices carried out by teachers in each profile. There were differences between the teachers of the competence profile and the rest, as shown in Table 3. Their assessment practices showed a continuous and practical approach, highlighted by the use of practical exams and a greater use of classroom practices. In addition, it was common for teachers of this profile to evaluate attendance systematically.

**Table 3.**

*Assessment practices in each profile.*

	N	FE	PrE	PE	PO	AS	PR	AT
Classic	7	5	0	3	2	3	3	2
Competence	5	3	3	3	0	3	5	4
Cohesive	4	3	0	0	0	0	3	1

Note: FE = Final Examination, PrE=Practical examination, PE=Partial Examination, PO=Portfolio, AS=Assignments, PR=Practices, AT=Attendance

**RQ2 - Are there differential characteristics in the teachers belonging to each of the profiles?**

### **Assessment training**

This research question searched for differences in the academic domain, teaching experience, in-service training, and assessment conceptions and practices among the teachers in each profile. In relation to the academic domain, there were no

appreciable differences, with representatives of the three profiles in each of the degrees. There were differences, however, in the teaching experience of each profile, as shown in Table 4. Specifically, the teachers categorised in the competence profile seemed to be those who had less experience, for both in teaching in higher education and teaching in their current subjects.

**Table 4.**

*Degree and experience of teachers in each profile.*

	N	Teaching experience (years)	Current subject experience (years)
Classic	7	22.3 (SD = 7.52)	7.1 (SD = 6.12)
Competence	5	13.8 (SD = 8.56)	4.6 (SD = 2.30)
Cohesive	4	21 (SD = 13.11)	10.3 (SD = 10.14)

Regarding in-service training, competence teachers were also the ones who did more teacher training courses, despite being the less experienced teachers. As shown in Table 5, teachers from this profile had done a total of 14 courses, mostly related to teaching and assessment practices.

**Table 5.**

*Teaching training courses by teachers in each profile.*

Topic	Classic	Competence	Cohesive
Assessment methods and tools	3	3	3
Teaching methods	3	5	0
Innovation	1	2	2
Use of ITCs	1	1	1
Syllabus elaboration	0	2	0
Interaction with students	0	1	0
<b>Average per teacher</b>	<b>1,14</b>	<b>2,8</b>	<b>1,5</b>

### **Assessment conceptions and practices**

Teachers in the three profiles showed different trends in the way they have developed their assessment practices over time. The first question they were asked in

this regard is why they had opted for the methods they currently use in their subjects. The teachers of the classic profile did not offer a clear explanation, and their answers indicated a certain inertia with respect to what was the tradition in their department or their faculty.

*A no-exam subject... it would be a bit weird. The student congress [an assessment instrument consisting in students' presentations simulating a scientific conference] is basically inherited from an old subject of the previous plan.*

*Teacher 9 - Medicine*

Over the years, they have tried to give greater weight to continuous assessment, but without modifying the set of instruments they had established. Instead, they varied the weight of each instrument on the final grade to make students focus more on those related to continuous assessment.

The conceptions that these teachers had about assessment had also changed over time. They claimed to rely on the students' responses to check whether their assessment practices were working or not. These answers had led them to reflect on the meaning of the final exam, and to consider alternative assessment practices. However, they were reluctant to make meaningful modifications of their assessment practices, and most of them argued departmental requirements as a constraint.

*Regarding the assessment criteria, each teacher in his group has freedom. But that freedom exists up to a certain point (...) There are aspects in which the teacher cannot... I do not think he could say... "well, I'm not going to make exams".*

*Teacher 5 - Mathematics*

The teachers of the competence profile generally declared that their assessment practices were designed with the aim of preparing their students for their professional future, trying to give them a practical and competence approach. However, they stated that they were limited by the time that a formative and continuous assessment requires, which they considered sometimes overwhelming. Regarding their development over the years, these teachers had also focused on promoting continuous assessment, but unlike the classic profile, with much more freedom to modify their assessment practices. In relation to their conceptions of assessment, they showed a growing interest in preparing their students for their professional future, adopting a practical and competence approach. In addition, they were aware of the changes that student cohorts have undergone since they began teaching, and they tried to adapt to them.

*Our students' brains now work in a different way from when I was studying.*

*Because society is changing, and things must be immediate, they have to be a click away. And then I think that sometimes we change things so that they have it just one click away, so they actually read it. And of course, if it costs them more than that, we already limp.*

*Teacher 4 - Sport Sciences*

The teachers of the cohesive profile declared that the assessment of their subjects had been a logical result of their teaching methods, or the characteristics of the subject. These teachers did not offer a concrete answer about the development of their assessment practices over the years.

*I have a colleague with whom I have a good relationship. So, I knew that he did project-based learning. I have to tell you; at that time, I did not have much faith in it. Well, we talked, we had coffee and I asked him how he did it. So, well, I started going to his classes ... in fact I think that that course I went to many*

*classes to see how he did it ... And well, I saw the dynamics of that project-based learning, so that was one of the influences to teach using project-based learning.*  
*Teacher 2 - Sport Sciences*

**RQ3 – What are the differences between the assessment practices designed during the assignment and the assessment practices implemented in their subjects?**

Given that the set of contents and learning outcomes to be assessed during the task was the same that the teachers must assess in their subjects, similarities and differences between the assessment practices designed during the task and those of their subjects were analysed. However, 10 teachers (62.5%) implemented changes in their assessment practices during the task. According to the categorization of Bearman et al. (2017), seven teachers carried out “Essential design activities” and three teachers carried out “Selective design activities”. Regarding the first group, the changes proposed consisted of creating additional tasks (n = 4), eliminating existing tasks (n = 1), or changing their format (n = 2). The second group modified the percentages of the assessment tasks already implemented in their subject.

When asked about the reasons between the differences in the practices designed in the task and those of their subjects, most of their answers indicated that implementing these methods in real practice would be excessively difficult for several reasons. Among them would be the workload it would entail for them (n = 3) or for their students (n = 1), logistical or coordination difficulty (n = 2), and the ratio of students to teachers (n = 4). Interestingly, two teachers stated that the only reason they had not implemented these methods in their subject is because they had not thought about it.

*I think the time to do it has not yet come for me. But I do believe that it is being generated little by little ... because the assessment must be in tune with the teacher, and no matter how much they want to innovate, or want to do, if they do*

*not feel safe doing it, or have their own ... fears, and sometimes training, which we also sometimes lack regarding assessment.*

*Teacher 15 - Sport Sciences*

To summarise, this study presents three different profiles regarding assessment design, which have been called classic, competence, and cohesive. Each profile has differential characteristics in their approach to assessment design.

### **Discussion**

The aim of this study was to explore the way in which higher education teachers design their assessment practices. The results provide new information as they are based on simulated performance instead of self-reported data.

The first research question explored whether it was possible to identify different profiles in the design of assessment practices. The results show different approaches regarding assessment design, which have been called classic, competence, and cohesive profiles. It should be noted that there seems to be a certain consistency in the design of assessment tasks, as the participants designed all their assessment tasks following the same pattern.

The classic profile was the most common among the participants. These teachers do not make considerations prior to selecting the instrument to be used. Choosing a determined set of instruments is the first step in their assessment design process, ending later with stating the grading weight of each one. Few teachers in this profile make additional considerations.

It is remarkable that many of the teachers in this profile showed a more formative approach to assessment in the initial interview questions than in their simulation task responses. It may imply that, even if they acknowledge the value of formative assessment, they lack the abilities to translate it to their assessment practices.

This issue has been addressed in several studies (Norton et al. 2010; Smith, 2011), and a variety of reasons might explain it. Participants of this study argue that they are influenced by the traditional assessment practices in their faculty or department, a well-known issue regarding assessment design (MacDonald & Joughin, 2009). The lack of training and resources to change their assessment practices is also another possible explanation.

The predominance of this profile might be in line with results such as those of Ibarra-Sáiz & Rodríguez-Gómez (2010) or Panadero et al. (2019), which show a traditional approach to university assessment. If most of the teachers do not reflect on formative aspects during their assessment design, or lack the resources to put them in practice, it is to be expected that they perpetuate summative approaches, as has been the tradition until now.

The competence profile was next in use. The teachers in this profile design their assessment practices starting from the learning outcomes of the subject, reviewing them in detail and reflecting on what would be the best way to ensure that students have achieved them, while helping them to do so. These teachers also tend to reflect on the feedback offered and more innovative methodologies are proposed.

It is noteworthy that the teachers of this profile were also the least experienced, both in university teaching in general, and in their current subjects. The trend towards a more formative approach among the youngest teachers has been argued by studies such as by Quesada-Serra et al. (2016) and seems to be supported by our results. However, teachers in this profile were also the most trained ones, in terms of in-service training courses. It is argued that the importance of pre-service training programmes is to develop assessment skills among teachers (Picos & López Pastor, 2013). Our results

also show the need of designing in-service training courses oriented to providing higher education teachers with the ability to design a formative assessment methodology.

Teachers in the competence profile, however, also faced challenges regarding their assessment design. They declared having a hard time balancing their assessment practices, which are often very complex, with their available resources, in terms of workload and external help. Much has been studied regarding the efficacy of formative assessment in terms of student learning, but these results raise questions about its efficiency. Authors such as Higgins et al. (2010) detail several principles for increasing formative assessment efficiency, including strategic curriculum review, use of IT, self- and peer assessment, oral feedback, in-class assessment, and group assessment. Apart from the latter two, none of them were mentioned by the participants of this study.

Finally, the cohesive profile is based on coherence between teaching and assessment practices. Teachers in this profile were not limited to designing independent assessment practices, even if the simulation task pointed in that direction; rather, their design went further by including various aspects of the teaching-learning process. They usually started the process by raising the teaching methods, starting for example from project- or problem-based learning methodologies. Once this was done, their assessment practices were a natural consequence of the teaching methods employed.

Teachers in this profile showed a less structured thinking process regarding assessment practices themselves, as they spent most of the time outlining the subject in general. They were also the ones who had been teaching their current subjects the longest, which may imply that some experience and trial-and-error processes are necessary to get to this point. Previous research exploring thinking processes in novice and experienced teachers seems to support this idea. Hall & Smith (2006) argued that

experienced teachers use less structured and more implicit thinking processes when planning their subjects. It also seems to be the case in the present study.

The last research question explores the changes between the assessment practices designed during the task and the assessment practices implemented in the actual teachers' subjects. Since the task was a mere simulation, with no necessary consequences in practice, the comparison between the assessment practices designed during it and those implemented in their subjects is appropriate to explore further the limitations of the assessment intended by the teachers when it is carried out in actual practice.

### **Implications for professional practice**

It is noteworthy that most of the teachers designed different assessment practices compared with the ones they use in their subjects, even though the aim was to assess the same set of learning outcomes. Most of the changes consisted of broadening the assessment practices, adding new tasks, or modifying existing ones to make them more complex. The main reason why they decided not to use these methods in their classrooms refers to the lack of resources, or the workload that it would imply, both for their students and themselves. This discovery aligns with previous studies such as those by Postareff et al. (2012), Norton et al. (2013), or Quesada-Serra et al. (2016), where teacher testimony showed many more formative approaches than there were put in practice.

The discrepancies between intentions and practice go back a long way. Norton et al. (2010) and Smith (2011) already showed that, even if teachers are aware of the formative purposes of education and agree with them, they are usually not capable of putting them into practice. Most of the participants recognise that their assessment

practices are not ideal, but rather an approximation that is achievable through the available resources.

The definition of ‘resources’ also varied among the participants of this study depending on their profile. Classic-profile teachers seemed to lack the knowledge and ability to make meaningful innovations to their assessment practices in practice, and they tended to use assessment practices coherent with the tradition in their department. Competence-profile teachers might have had the knowledge, as they were usually well-trained regarding teaching and assessment. However, even if they considered themselves capable of designing an effective set of assessment practices, they claimed to be unable to balance them and make them efficient in terms of workload and coordination.

Both challenges mentioned by the participants are relatively well-known among the assessment literature. Based on previous theory and research work, we offer a series of recommendations for institutions and teachers that aims to help them face the challenges regarding assessment design.

### **Teacher training and organizational solutions**

In the country where our study was based (Spain), pre-service pedagogical training is not compulsory for higher education teachers. This results in many teachers starting their assessment design processes without any literacy or theoretical knowledge about how assessment must be done. This makes it especially important to offer high-quality in-service training, to provide them with the empirical and practical tools to improve their assessment practices.

The previous literature provides some instructions for in-service training about assessment. First, authors such as Brookhart (2011), based in the seminal ‘Standards’ framework (American Federation of Teachers, National Council on Measurement in

Education, & National Education Association, 1990) offers a comprehensive list of assessment-related competences that teachers must acquire. Some examples of these competences are understanding learning in the content area they teach, having a wide repertoire of assessment strategies, or providing effective and useful feedback. For more information, see Brookhart (2011. p.7).

Nevertheless, a set of training courses covering the competences proposed by Brookhart can still be unsuccessful if not implemented correctly. Xu & Brown (2016) carried out a systematic review on assessment literacy, concluding that teacher education must address four pertinent issues to be effective for teachers.

First, teachers must have a solid assessment literacy knowledge base, implying that teacher education programs must include assessment courses as part of their curriculum (Sato et al., 2008). Second, assessment education needs to be long enough to let teachers acquire a deep understanding about assessment. It also needs to establish connections between assessment theory and practice (Lyon, 2013). Third, it needs to address teachers as individuals and professionals (Hill et al., 2010). Every teacher has a background of conceptions, expectations, and prior experiences about assessment. Lastly, teacher educators need to understand that assessment literacy development is not merely an accumulation of assessment knowledge, but rather the development of a sophisticated, contextually appropriate set of inter-related competencies (Xu & Brown, 2016).

It is also important to consider that in-service teachers have limited time and opportunities for pedagogical training, so institutions must adapt to these circumstances. On-line learning (Fan, Wang & Wang, 2011) and reflecting on their assessment practices (Smith, 2011) can be methods for them to learn without compromising their workload.

**Time management, individual solutions**

The other big issue mentioned by our participants was the lack of time to put in practice their preferred assessment methods. Efficiency in formative assessment has been an underexplored topic in the literature, but authors such as Higgins et al. (2010, p. 10) provide useful strategies for carrying out high-quality assessment methods without compromising the time available to the teacher. These strategies are as follows:

1. Strategic curriculum review: The assessment workload can be reduced by avoiding repetition. Continuous assessment is a requirement in many European universities, but there are methods to minimize the total number of products assessed. These methods could include exempting students from a final assessment based on coursework performance (Hornby, 2005, p.22).
2. Use of ICT: Using automatically assessed virtual tools is a way to obtain general knowledge of your students' progress without having to read and assess individual tasks.
3. Group assessment: It can mean less workload for the teacher while different skills are developed by the students (Higgins et al. 2010).
4. In-class assessment: can include possible periodic tests, assessing each other, and gaining exemption from the final exam (Hornby, 2005, p.23).

Additionally, oral informal feedback given during in-class activities can be easily combined with the use of ICT, and it is especially useful in large classes.

5. Peer and self-assessment: its impact on learning and self-regulation is one of the main lines of work on educational research these days (Panadero et al., 2016). It is also easily combined with the use of all the previous categories.

Apart from a method to reduce teacher workload, its use can be enormously beneficial for students, if implemented correctly.

### **Implications for future research**

These results provide a detailed view of what is happening when assessment is designed. Until now, research had focused on the factors, both personal and contextual, that influenced the process (McMillan, 2003; Bearman et al. 2017). However, despite the broad conceptual work published on this topic, there were few empirical studies exploring how this process is carried out in practice. The discovery of three different profiles is a great leap in research on assessment design in higher education. These results open a new avenue of research, with relevant questions for research that must be answered in subsequent studies.

The first question is if it is possible to identify predictors of the different profiles. One aim of the present study was to identify such predictors, and in our results, it was shown how teaching experience and in-service training may influence how teachers design their assessment practices. However, our small sample made it difficult to make solid claims regarding predictors.

Other factors such as disciplinary approaches, departmental culture, and student characteristics have been found to influence teachers' assessment design (Meyer et al. 2010; Bearman et al. 2017; Fernández Ruiz et al., under review). These variables could theoretically work as a predictor for the profiles found in this study. To find these predictors and to connect them with the profiles described here would be fundamental. This knowledge would allow teacher training programmes adapted to the personal necessities and approaches of each teacher and based on their real practices regarding assessment design, which is still a challenge for educators (Korthagen, Loughran & Russell, 2006).

The second question is about whether these approaches to assessment design remain stable throughout the teacher's professional career. Our results show a difference in the teaching experience among teachers in each profile. Less experienced teachers tend to show a formative approach in assessment design, and more experienced teachers are more inclined towards the competence and cohesive profiles. It is important to understand the cause of these differences.

It may be due to the new professional context, which would imply a heavier cognitive load as much of the processes have not been automated yet. In this case, once they have connected specific learning outcomes to specific teaching and assessment practices, they would focus on different factors during further redesigns, as was observed in the other profiles. This might be counterproductive, considering that both the learning outcomes and the social and professional environment can be enormously variable throughout a teacher's professional career. Teachers would then have difficulties to adapt their instructional and assessment practices to the current context.

On the contrary, the differences among teachers regarding experience can be an effect of a stronger scientific and institutional focus on formative assessment in recent years. In this case, these differences would be considered a positive outcome of this new approach and it would be expected that teachers maintain the same approach throughout their professional careers.

The third question is about the possibility of combining several profiles. The profiles identified in this study show three different approaches to assessment design. Teachers seem to focus on available resources, learning outcomes, or teaching method alignment, and the combination of these is scarce. However, the three approaches are fundamental to achieve high-quality assessment practices. It is desirable that university teachers are able to design an assessment methodology aligned with their teaching

methods and the learning outcomes, but also feasible and efficient considering their available resources. It is necessary to further explore which would be the best way to help the teachers to balance their assessment designs paying attention to the three factors.

### **Limitations**

As the present study had a comprehensive and in-depth approach to data, the number of participants was smaller than survey type studies. Having a sample consisting of only 16 teachers may affect the significance of the findings and its applicability in other contexts. A significant effort was made to maintain a diverse sample in terms of institution (four universities), discipline, and teaching experience. Also, teachers from our study reported assessment design processes highly related with the constraints they face in their practice. Exploring the applicability of our results to teachers from different professional backgrounds must be done in further studies.

### **Conclusion**

The present study shows an innovative research approach to investigate how university teachers design their assessment practices. We found three different design patterns: classic, which focuses on the feasibility of the assessment tasks; competence, which focuses on alignment with the learning outcomes; and cohesive, which focuses on alignment with their teaching methods. It is striking that most of our participants carried out a design focused only in one of the three elements. Interestingly, the assessment practices that the participants designed in the simulation task were different from the ones they use in their real lectures. When asked about this point, participants explained these differences because they lack the resources needed to implement these complex practices. This study represents a starting point for practice-based research involving assessment in higher education. Our results make explicit the need to answer

further questions to fully understand how higher education teachers can be helped to link their assessment practices with the vast empirical knowledge about assessment.

### **Disclosure statement**

No potential conflict of interest was reported by the authors.

### **References**

- American Federation of Teachers, National Council on Measurement in Education, & National Education Association (AFT, NCME, & NEA). (1990). The standards for teacher competence in the educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30–32.
- Andrade, H. L., & Heritage, M. (2017). *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Routledge.
- Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E., Boud, D., & Joughin, G. (2017). How university teachers design assessments: a cross-disciplinary study. *Higher Education*, 74(1), 49-64.
- Bennett, S., Thomas, L., Agostinho, S., Lockyer, L., Jones, J., & Harper, B. (2011). Understanding the design context for Australian university teachers: Implications for the future of learning design. *Learning, Media and Technology*, 36(2), 151–167. doi:[10.1080/17439884.2011.553622](https://doi.org/10.1080/17439884.2011.553622).
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & evaluation in higher education*, 31(4), 399-413.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: issues and practice*, 30(1), 3-12.

- Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education, 69*(6), 963-976.
- Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The clearing house: A journal of educational strategies, issues and ideas, 83*(1), 1-6.
- Fan, Y. C., Wang, T. H., & Wang, K. H. (2011). A web-based model for developing assessment literacy of secondary in-service teachers. *Computers & Education, 57*(2), 1727-1740.
- Fernández Ruiz, J., & Panadero, E. (2020). Comparison between conceptions and assessment practices among secondary education teachers: more differences than similarities. *Infancia y Aprendizaje: Journal for the Study of Education and Development, 43*(2), 309-346.
- Fernández Ruiz, J., Panadero, E. & García-Pérez, D. (2021). Assessment from a disciplinary approach: Design and implementation in three undergraduate programmes. Manuscript submitted for publication.
- Hall, T. J., & Smith, M. A. (2006). Teacher planning, instruction and reflection: what we know about teacher cognitive processes. *Quest, 58*(4), 424-442.
- Higgins, M., Grant, F., & Thompson, P. (2010). Formative assessment: balancing educational effectiveness and resource efficiency. *Journal for Education in the Built Environment, 5*(2), 4-24.
- Hill, M., Cowie, B., Gilmore, A., & Smith, L. F. (2010). Preparing assessment-capable teachers: What should preservice teachers know and be able to do?. *Assessment Matters, 2*, 43-64.
- Hornby, W. (2005). Dogs, stars, Rolls Royces and old double-decker buses: Efficiency and effectiveness in assessment. In: *Enhancing practice: Reflections on*

- assessment: Volume 1*, 15–28. Gloucester: Quality Assurance Agency for Higher Education.
- Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2010). An approach to the dominant discourse of learning assessment in higher education. *Revista de Educación*, 351, 385-407.
- Izci, K., & Caliskan, G. (2017). Development of Prospective Teachers' Conceptions of Assessment and Choices of Assessment Tasks. *International Journal of Research in Education and Science*, 3(2), 464-474.
- Jessop, T., & Tomas, C. (2017). The implications of programme assessment patterns for student learning. *Assessment & Evaluation in Higher Education*, 42(6), 990-999.
- Korthagen, F., Loughran, J., & Russell, T. (2006). Developing fundamental principles for teacher education programs and practices. *Teaching and teacher education*, 22(8), 1020-1041.
- Lyon, E. G. (2013). Conceptualizing and exemplifying science teachers' assessment expertise. *International Journal of Science Education*, 35(7), 1208-1229.
- Macdonald, R., & Joughin, G. (2009). Changing assessment in higher education: A model in support of institution-wide improvement. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 1–21). Rotterdam, Netherlands: Springer.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational measurement: Issues and practice*, 22(4), 34-43.
- McMillan, J. H., & Moore, S. (2020). Better being wrong (sometimes): classroom assessment that enhances student learning and motivation. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 93(2), 85-92.

- Meyer, L. H., Davidson, S., McKenzie, L., Rees, M., Anderson, H., Fletcher, R., et al. (2010). An investigation of tertiary assessment policy and practice: Alignment and contradictions. *Higher Education Quarterly*, *64*(3), 331–350.  
doi:[10.1111/j.1468-2273.2010.00459.x](https://doi.org/10.1111/j.1468-2273.2010.00459.x).
- Norton, L., Aiyegbayo, O., Harrington, K., Elander, J., & Reddy, P. (2010). New lecturers' beliefs about learning, teaching and assessment in higher education: the role of the PGCLTHE programme. *Innovations in Education and Teaching International*, *47*(4), 345-356.
- Norton, L., Norton, B., & Shannon, L. (2013). Revitalising assessment design: what is holding new lecturers back?. *Higher Education*, *66*(2), 233-251.
- Norton, L., Richardson, T. E., Hartley, J., Newstead, S., & Mayes, J. (2005). Teachers' beliefs and intentions concerning teaching in higher education. *Higher education*, *50*(4), 537-571.
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 133-148.
- Panadero, E., Brown, G., & 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.
- Panadero, E., Brown, G. T., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, *28*(4), 803-830.

- Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, 44(3), 379-397.
- Picos, A. P., & 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. Do What I Say, Not What I Do: Student Assessment Systems. *Revista de educación*, 361, 279-305.
- Postareff, L., Virtanen, V., Katajavuori, N., & Lindblom-Ylänne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation*, 38(3-4), 84-92.
- 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.
- Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal*, 45(3), 669-700.
- Smith, J. (2011). Beyond evaluative studies: perceptions of teaching qualifications from probationary lecturers in the UK. *International Journal for Academic Development*, 16(1), 71-81.
- Stiggins, R. (2007). Assessment through the student's eyes. *Educational leadership*, 64(8), 22.
- Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science*, 33(4), 283-309.

- Wu, Q., & Jessop, T. (2018). Formative assessment: missing in action in both research-intensive and teaching focused universities?. *Assessment & Evaluation in Higher Education*, 43(7), 1019-1031.
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162.