Contextual Contrast of Assessment Practices in Catalonia and Tuscany

David Martínez-Maireles[*], Davide Capperucci[**], Mila Naranjo Llanos[***]

1. Introduction

Government proposals for change and innovation projects in the field of education generate anxiety and uncertainty in teachers, since they are accompanied by initiatives or specific models to be developed. and they impose regulation, bureaucracy and organisational rigidity. This can create a misalignment between educational policies and teachers' practice that makes it difficult to implement innovation (Caliskan & Zhu, 2019).

A more inclusive policy of educational practices is promoted both in Catalonia[1] (Spain) and Italy[2], which in turn affects the assessment practices. Italian laws have been focusing on inclusion and alignment of the teaching-learning-assessment process for a longer period. Their regulations are more prescriptive than Catalan laws on the object of assessment and the development of assessment practices, although the regulation of educational competencies that must be promoted in educational processes is better defined in

Catalonia than in Italy. This necessary coherence between all the elements of the teaching-learning-assessment process allows for an alignment that serves both assessment of learning process by teachers and the introduction of modifications and improvements in teaching (Ciani, Ferrari, & Vannini, 2020).

The aim of this article is to propose a contextual comparison between the regions of Catalonia and Tuscany focused on teachers' perception of the importance, competence and use of the planning and development of the assessment practices, which includes their improvement and monitoring of students' learning and participation. The study focuses on pre-school, primary and secondary school education.

2. Theoretical framework

The theoretical framework is divided into four sections: planning the assessment, monitoring students' learning, encouraging student participation in the assessment, reviewing, improving and innovating assessment. The same sections that make up the instrument were used to carry out the study of the assessment practices of the teachers.

2.1. Planning the assessment

According to OFSTED (2008), assessment must be a central element in the planning of the curriculum, especially in the definition of competencies and learning objectives, in order for it to have a positive impact.

If the assessment is not aligned, that is, if it lacks any connection with knowledge, competencies and learning objectives, it will have neither continuity nor coherence with teaching and learning processes, and this will affect its validity (Cizek, Kosh, & Toutkoushian, 2018).

Consequently, two aspects are important: planning the assessment practices in advance in order to adjust and align them with the teaching-learning process and ensuring that they are co-constructed by the teachers and also with the students so as to increase their involvement and their self-regulation (Folch, Córdoba, & Ribalta, 2020). Once planned, the assessment must be carried out incorporating important aspects involving monitoring of the students' learning, their participation and improvement of the assessment practices.

34

2.2. Monitoring students' learning

If the planning of the assessment practices is carried out as indicated, in order to adjust them to the teaching-learning process of the teachers and the students, during their development *the students' learning must be monitored* for three moments and types of assessment: initial assessment (diagnostic), continuous assessment (formative) and final assessment (summative) (Black & Williams, 2018).

First, the assessment needs to consider previous student knowledge. For this, Murillo & Duk (2015) and Vannini (2011) propose carrying out a diagnostic assessment in order to gather information about the level of learning and the educational and formative needs of the students at the outset, with a triple purpose: define learning objectives, generate programming based on these objectives, and make students aware of the nature of the teaching-learning process.

Second, it is essential to use formative assessment for the teaching-learning-assessment process to promote student autonomy and self-responsibility (Vannini, 2011; Trinchero, 2017). An assessment that provides essential information on the quality of the teaching and learning praxis for subsequent feedback and feedforward (Cornoldi, De Beni, & Gruppo, 2020; Li & Grion, 2019). This information increases the responsibility and autonomy of the students regarding their own process, which will be helpful to them in their later learning (Analí, Paoloni, & Donolo, 2017). Third, if *summative assessment* serves not only the purpose of accountability but is also accompanied by formative feedback linking it to previous assessments, the whole assessment process will be at the service of the student's self-regulation and autonomy (Black & Williams, 2018; Darmody, Lysaght, & O'Leary, 2020; Koenen, Dochy, & Berghmans, 2015; Kulamankan & Rangachari, 2018).

2.3. Encouraging Student Participation in the Assessment

The need to increase the degree of *student participation in assessment* practices is thus clear, since it improves students' performance, competencies, motivation, and the capacity to self-organize their learning (Cornoldi, De Beni, & Gruppo, 2020; Li & Grion, 2019; Lucisano & Stanzione, 2018). In addition, it enables them to become more aware of and responsible for their learning processes and make the assessment more effective if the assessment criteria are transformed into learning criteria and are shared with the students, and heterogeneous assessment, self-assessment and co-assessment are promoted (Murillo & Duk, 2015). This also increases autonomy,

promotes individual and team metacognition and the transference of competencies (Adachi, Hong-Meng Tai, & Dawson, 2017).

This focus on the students, increasing their involvement in the assessment practices through joint decision-making, choosing the learning objectives to be assessed, establishing and applying correction criteria in assessment tasks, also supports an increase in the self-regulation of their own process (Koenen, Dochy, & Berghmans, 2015; Kulamankan & Rangachari, 2018).

Choosing to foster student participation and to monitor their learning process also makes it necessary to review the assessment practices carried out and to introduce improvements to adjust and align the whole teaching-learning-assessment process.

2.4. Reviewing. improving and innovating assessment

This review, *improvement and innovation of the assessment practices* to ensure their constant development needs to be promoted by the institution itself by offering and promoting opportunities for exchange and joint reflection among teachers (Koenen, Dochy, & Berghmans, 2015). This would increase the perception of competence in coordination and collaboration by teachers, shown to be very low in the study by Quesada, Rodríguez, & Ibarra (2017), since sharing their practices with their colleagues and opening up to other teaching options not only promotes cooperative and collaborative work between them inside and outside the classroom, which can benefit the students, but also promotes the personal and professional development of teachers (Folch, Córdoba, & Ribalta, 2020; King-Sears & Strogilos, 2020).

Teachers attach great importance to assessment practices, although it is necessary to update teacher training related to these practices in order to extend their use and alignment with the teaching-learning process, and more so if improvements are made in the latter (Caliskan & Zhu, 2019; Quesada, Rodríguez, & Ibarra, 2017).

Moreover, it is necessary that training in assessment practices and innovation processes in this area should be promoted by the government and the management team, and that the latter needs to recognize its importance and become involved in the process by creating the time and organizational structure necessary to carry them out, since otherwise it will be difficult for them to be implemented and for continuity to be assured (Solheim, Roland, & Ertesvåg, 2018).

It is important to link these four dimensions of assessment practices – *assessment design and planning*; *monitoring of student learning*; *promotion of student participation in the assessment practice*; and *reviewing, improving*

36

and innovating assessment – with three criteria: their *importance, perceived competence* and *use* by the teachers (Flores, 2012; Kraft & Gilmour, 2016; Jiang, Sporte, & Luppescu, 2015). This is important because knowing teachers' opinions regarding which aspects of these four dimensions they consider more or less important, in which they perceive themselves as more or less competent, and which they use to a greater or lesser extent, helps and has an impact on two levels. First, it enables teachers to become aware of their own practice. Second, it helps the management team develop a general vision and propose improvements to the school's assessment policy that are more aligned and adjusted to the reality and needs of the professionals.

3. Methods and materials

In order to achieve a contextual comparison of assessment practices in Catalonia (Spain) and Tuscany (Italy) a descriptive survey study using a questionnaire was carried out.

3.1. Participants

173 teachers from two countries (7 schools in Catalonia; 4 schools in Tuscany) took part in the study.

Table 1 – *Distribution of participating teachers by school stage and region*

	Pre-school (3-6 years)	Primary school (6-12 years)	Compulsory secondary education (12-16 years)/ 1st grade Tuscany (12-14 years)	Total
Catalonia	3	13	8	24
Tuscany	17	78	54	149

3.2. Data collection instrument

In order to carry out the contextual comparison of the level of *importance, perceived competence* and *use* of the assessment practices in Tuscany and Catalonia, the ActEval questionnaire on assessment practices (Quesada, Rodríguez, & Ibarra, 2013) was chosen. This questionnaire is divided into four dimensions and each dimension consists of a different number of items that add up to a total of 31. Each item uses a Likert scale (1 = never; 6 = completely) for three different criteria; the importance that the teacher attaches

37

to this item; their *perceived competence*; and the use they make of the item in question.

This instrument was chosen to carry out the research with teachers of pre-school, primary and high school, despite being aimed at university teachers, because in the study of the three criteria of the assessment practices there are still no specific instruments that evaluate the assessment practices in education stages below university level, where it is most studied. The questionnaire was subjected to a study of face and content validity, with a resulting Chronbach's alpha of 0.9.

3.3. Procedure

First, we contacted the school principals of the different schools in Catalonia, who chose the teachers that would participate in the study. In Tuscany, we sent an invitation to which four schools responded, and all the teachers of the schools were invited to participate voluntarily. Subsequently, a meeting was held with the school principals of the school to explain the project and what their participation would involve.

The link to the anonymous questionnaire was then sent, with a response period of one month that had to be extended to three months due to Covid-19. The questionnaires were collected, the responses collated, and a descriptive analysis (mean, standard deviation and frequency) carried out with SPSS software and completed with a parametric analysis with Student's T-test of independent groups with a significance of $p < .05$, after categorizing the data, to check whether there were any statistically significant differences between the two regions. This test rather than a non-parametric test was chosen because the sample from Tuscany was over 30 (N = 149), although the sample from Catalonia was smaller (N = 24).

4. Results

This section presents the results of the questionnaire divided into three parts. Firstly, the frequency tables of each criterion – *importance* I, *competence* C and *use* U – for each item in each dimension and the mean graphs of each criterion for each item in each dimension in Tuscany. Secondly, the frequency tables and the mean graphs in Catalonia. Thirdly, the Student t-test of each criterion for each item in each dimension in the two regions.

4.1. Tuscany

4.1.1. Dimension 1: assessment design and planning

Table 2 – *Frequency as a percentage of the importance, competence and use criterion for each item of the assessment design and planning dimension in Tuscany*

| Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
1	73.8	56.4	56.3	22.4	38.9	38.3	2.6	4.7	5.4
8	76.5	60.4	55.0	19.5	35.6	39.6	4.1	4.0	5.3
9	73.8	59.0	53.7	21.5	36.2	40.3	4.7	4.7	6.1
15	73.2	59.1	50.3	22.9	33.6	41.6	4.0	7.4	8.1
18	70.4	46.9	44.3	26.1	61.8	51.0	3.3	5.4	4.7
19	57.8	40.3	36.9	36.9	49.0	48.3	5.4	10.7	14.8
20	77.2	64.4	63.8	20.8	32.9	33.6	2.0	2.6	2.6
25	78.5	73.8	69.1	19.4	23.5	26.2	2.0	2.7	4.7
31	36.9	26.2	20.8	43.0	36.9	35.6	20.2	36.9	43.6
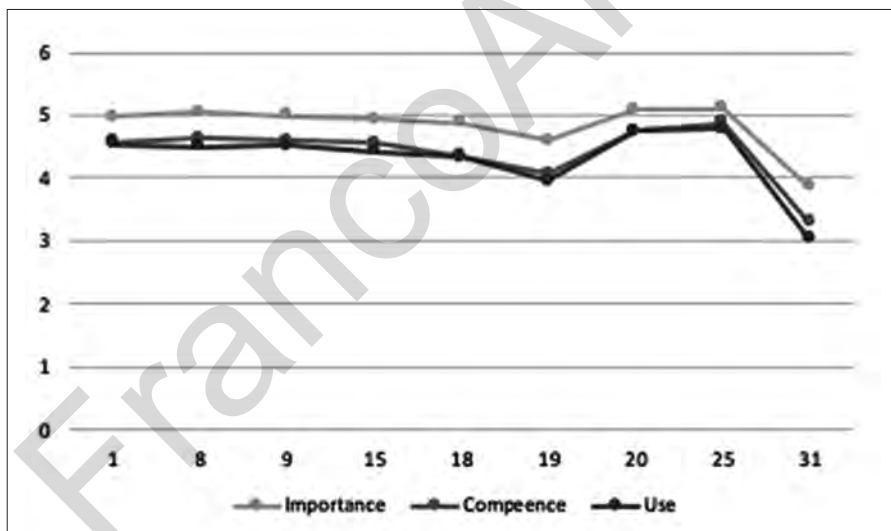
Figure 1 – *Mean per criterion for each item of the assessment design and planning dimension in Tuscany*

As can be seen in the results for *assessment design and planning* dimension, the criteria have a relatively even distribution, particularly the two criteria of perceived *competence* and *use*, whose frequencies and mean are only a few

39

points apart. Nevertheless, the teachers see the adaptation to use of electronic media for teaching (19) as very *important* and consider themselves quite *competent*, which has an impact on their *use*. However, the results differ on the use of technological platforms and tools in the assessment (31), which is perceived as fairly or not very *competent* and a higher number use it little or never.

This raises the question of whether there is a complete alignment between the educational process, as shown by the results of item 20 on considering assessment in an integrated way in the process.

4.1.2. Dimension 2: monitoring student learning

Table 3 – *Frequency as a percentage of the importance, competence and use criterion for each item of the monitoring of student learning dimension in Tuscany*

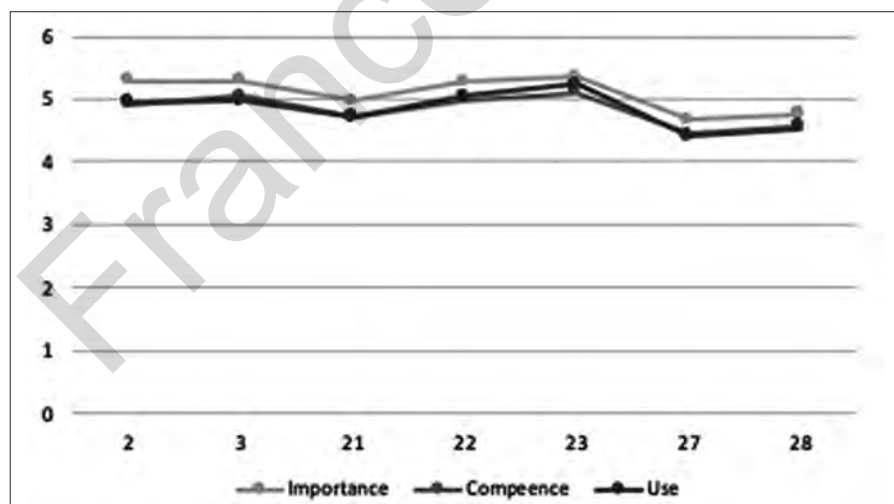| Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
2	83.9	73.8	75.2	14.8	23.5	22.1	1.4	2.6	2.6
3	82.5	77.2	73.8	16.7	21.5	24.8	0.7	1.4	1.4
21	70.5	63.2	61.8	26.8	33.6	34.2	2.7	3.3	4.0
22	81.2	73.8	77.1	18.1	24.2	20.2	0.7	2.0	2.6
23	81.2	79.2	83.9	18.1	19.4	14.8	0.7	1.3	1.4
27	60.4	53.1	51.7	34.9	42.9	40.3	4.7	4.0	8.1
28	66.5	60.4	59.1	29.6	32.9	33.5	6.0	6.7	7.4

Figure 2 – *Main per criterion for each item of the monitoring of student learning dimension in Tuscany*

40

Compared to the other dimensions, the results for *monitoring student learning* dimension show that teachers' perceived *competence* and *use* are closer to importance. It should be noted that teachers give high scores to all three criteria in terms of providing feedback and feedforward on results so that students can modify and improve their performance (2, 3). However, priority continues to be given to final assessment (23), rather than initial (21) and formative (22). Medium *importance* is given to using assessment to meet the needs of learners (27) and this means that its *competence* is also split between high and medium scores. This is perhaps due to a lack of training.

4.1.3. Dimension 3: promoting student participation in assessment

Table 4 – *Frequency as a percentage of the importance, competence and use criterion for each item of the promotion of student participation in assessment dimension in Tuscany*

Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
4	81.9	64.4	63.7	15.5	30.9	31.5	2.7	4.7	4.7
10	73.1	59.8	54.4	23.5	36.3	40.9	3.3	4.0	4.7
11	69.8	53.7	53	25.5	41.6	40.2	4.7	4.7	6.7
12	75.9	65.7	58.4	20.8	29.5	34.9	3.4	4.7	6.7
13	63.8	53.7	44.3	26.8	36.2	37.6	9.4	10.1	18.1
14	50.4	38.9	33.0	36.9	43.7	45.6	12.8	17.5	21.5
16	71.1	64.5	59.0	23.5	29.5	31.5	5.4	6.1	9.4
17	70.5	61.7	54.4	26.9	33.2	39.6	2.7	4.0	6.1
26	67.1	59.0	51.0	19.5	37.6	41.0	3.4	4.7	8.1
29	61.7	47.6	41.6	32.2	42.3	44.9	6.0	13.0	13.4

In *promotion of student participation in assessment* dimension, the items that obtained the highest scores were those referring to four aspects: awareness of their role in assessment practices (10), reaching a consensus on the object and criteria of assessment (16), the mechanics of assessment (4), agreeing and participation through self-assessment (12).

Although, apart from self-assessment, peer assessment (13) and co-assessment (14) could also be used, these two types of assessment received the lowest scores.

Compared to the other dimensions, this is the dimension where use is least aligned with perceived competence, although there is not a large gap between them.
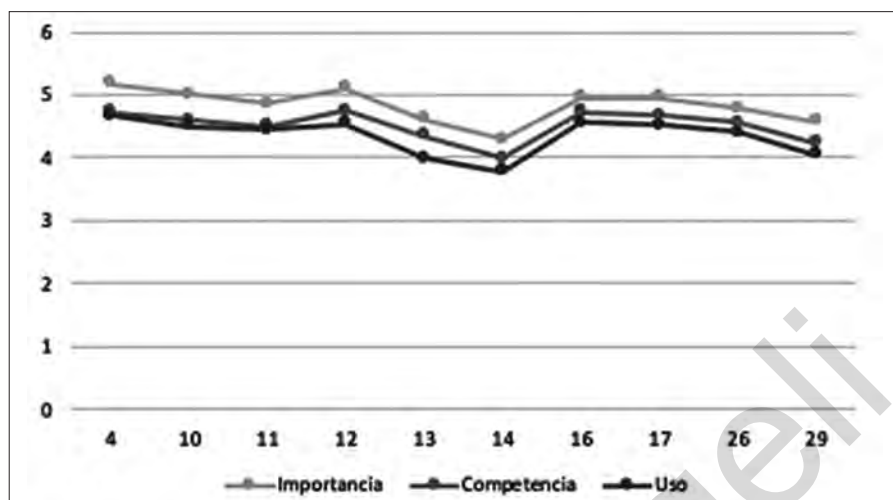
Figure 3 – *Mean per criterion for each item of the promotion of student participation in assessment dimension in Tuscany*

4.1.4. Dimension 4: reviewing. improving and innovating assessment

Table 5 – *Frequency as a percentage of the importance. competence and use criterion for each item of the reviewing. improving and innovating assessment dimension in Tuscany*

Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
5	77.2	57.1	53.7	20.8	40.2	42.3	2.0	2.7	4.1
6	66.4	49.0	45.6	30.2	45.7	46.3	3.4	5.4	8.0
7	80.5	59.0	55.0	16.1	36.2	37.6	3.4	4.7	7.4
24	75.1	63.8	61.1	21.5	30.9	34.2	3.4	5.4	4.7
30	75.2	59.8	51.0	21.5	35.6	40.2	3.4	4.7	8.8

In the results for *reviewing, improving and innovating in assessment* dimension, it should be noted that teachers update their knowledge of learning assessment (7), but when it comes to making improvements (5, 6), they are divided between those who perceive themselves to be highly competent and use it, and those who perceive themselves to be at a medium level.

Item 30, favoring collaboration and coordination between teachers in assessment processes, is considered by teachers to be very *important*. However, in its use, teachers are divided between those who use it more often and those who use it at a medium level.
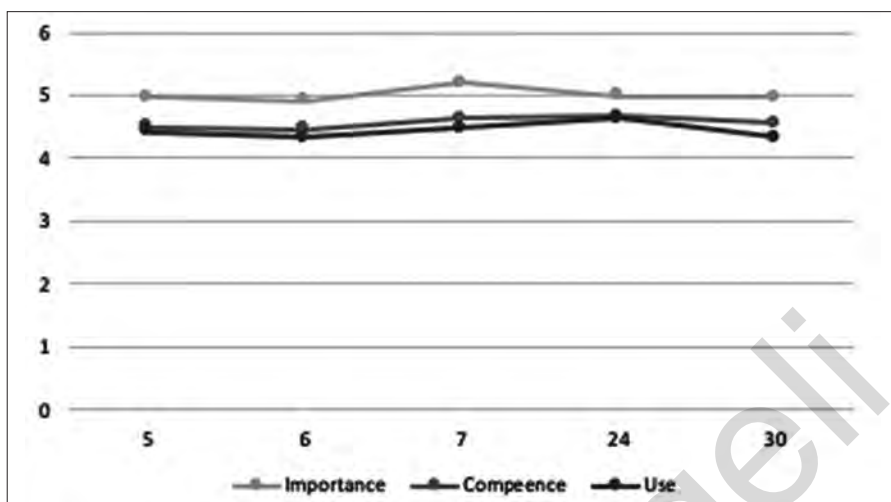
42

Figure 4 – *Mean per criterion for each item of the reviewing, improving and innovating assessment dimension in Tuscany*

4.2. Catalonia

4.2.1. Dimension 1: assessment design and planning

Table 6 – *Frequency as a percentage of the importance, competence and use criterion for each item of the assessment design and planning dimension in Catalonia*

Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
1	87.5	39.1	60.0	8.3	52.0	28.0	8.7	4.7	5.4
8	41.7	13.0	28.0	33.3	43.4	32.0	26.1	4.0	5.3
9	100.0	65.2	68.0	–	34.8	28.0	–	4.7	6.1
15	91.6	39.1	48.0	8.4	47.8	36.0	13.0	7.4	8.1
18	83.4	39.1	28.0	12.5	43.4	48.0	8.7	5.4	4.7
19	16.7	13.5	12.0	54.2	43.4	16.0	26.1	10.7	14.8
20	93.3	69.6	60.0	4.2	21.7	28.0	–	2.6	2.6
25	79.2	47.8	52.0	16.7	43.4	40.0	–	2.7	4.7
31	41.6	26.0	24.0	45.9	43.4	24.0	21.7	36.9	43.6

In the results for *assessment design and planning* dimension, a rather uneven distribution can be observed. Item 19 (adaptation of the teaching process for the use of electronic media) stands out as being perceived as having a low *importance*, which has an impact on its *competence* and *use*. This is

43

linked to the use of technological tools in the assessment process (31), since if there is no adaptation to these platforms it is difficult to have a high use of them in assessment.
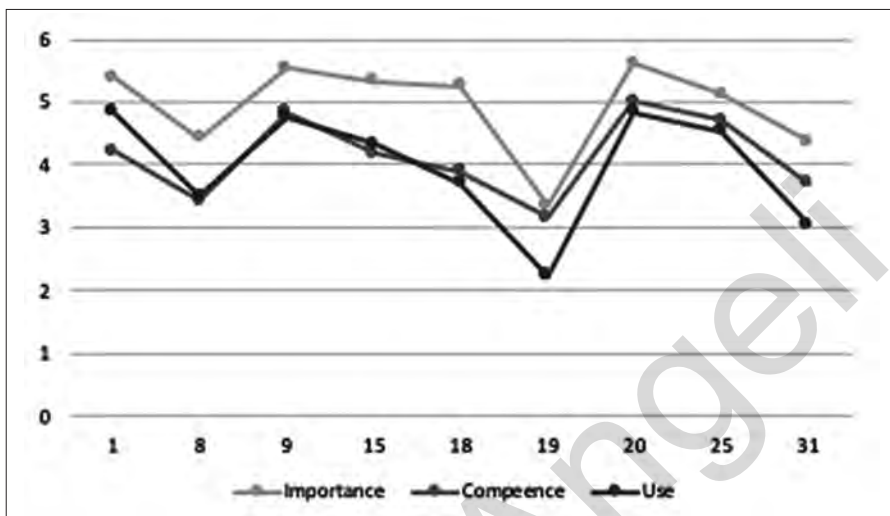

Figure 5 – *Mean per criterion for each item of the assessment design and planning dimension in Catalonia*

As regards the design of assessment systems and processes and the construction of objectives and instruments (1, 15, 18), it can be seen that great importance is attached to these, but their competence and use decreases the more concrete the aspect to be designed is. In addition, make the objectives, standards and criteria of education known (19) is more *important*, they feel more *competent* and they *use* it more than make the qualification procedure known (25), perhaps because there is more of a culture of sharing the aspects of the teaching process than those of assessment.

4.2.2. Dimension 2: monitoring student learning

In the results for *monitoring of student learning* dimension, it is worth noting the high score on the *importance* criterion for all items. There is also an alignment and a high score on the *competence* and *use* of the three types of assessment: initial (21), formative (22) and final (23). However, aspects that would support them, such as feedback (2) or feedforward (3), and promote learning through it (28) are rated as medium *competence* and *use*, perhaps due to lack of training.

44

Table 7 – *Frequency as a percentage of the importance, competence and use criterion for each item of the monitoring of student learning dimension in Catalonia*

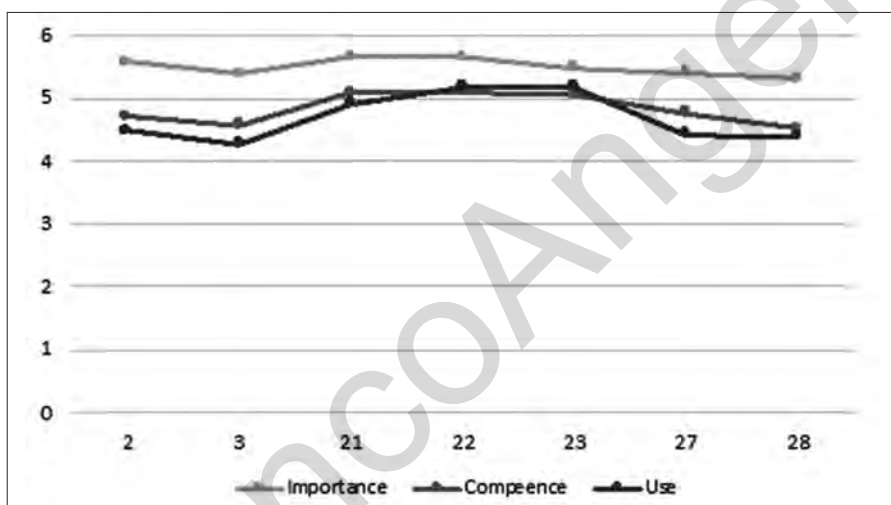| Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
2	95.8	56.5	48.0	4.2	43.4	52.0	–	–	–
3	91.7	56.5	40.0	4.2	43.5	60.0	–	–	–
21	87.5	78.2	68.0	8.3	13.0	24.0	–	4.3	4.0
22	91.7	73.9	76.0	4.2	13.0	8.0	–	–	4.0
23	91.7	73.9	76.0	4.2	21.7	20.0	–	–	–
27	87.5	56.5	44.0	4.2	30.4	40.0	–	4.3	8.0
28	79.2	43.5	40.0	12.5	47.8	52.0	–	–	–

Figure 6 – *Main per criterion for each item of the monitoring of student learning dimension in Catalonia*

4.2.3. Dimension 3: promoting student participation in assessment

The results for *promotion of student participation in assessment* dimension show that the competence and use of student participation in assessment design (11) has a medium-low score, although a consensus is promoted with students on what is assessed (16) and with what criteria (17), and to a lesser extent on the grading procedure (26). Their participation is mainly promoted through self-assessment (12), peer-assessment (13) and co-assessment (14). This gives meaning to the score for item 10 (make students aware of the benefits of their participation in the assessment practice) which is the highest in all three criteria.

Table 8 – *Frequency as a percentage of the importance, competence and use crite-rion for each item of the promotion of student participation in assessment dimension in Catalonia*

| Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
4	70.8	30.4	36.0	20.8	56.5	48.0	8.4	13.0	16.0
10	91.6	52.2	56.0	8.4	47.8	36.0	–	–	8.0
11	62.5	34.8	20.0	29.2	52.2	44.0	4.2	8.7	32.0
12	95.8	52.1	52.0	4.2	43.5	44.0	–	4.3	4.0
13	87.5	52.1	48.0	12.5	43.5	48.0	–	4.3	4.0
14	70.8	43.5	32.0	29.2	43.4	48.0	–	13.0	20.0
16	65.8	52.1	48.0	29.2	39.1	40.0	–	–	8.0
17	75.0	52.2	32.0	20.8	43.4	52.0	–	–	8.0
26	45.8	34.8	16.0	41.7	43.4	60.0	8.3	17.4	20.0
29	54.2	39.1	24.0	37.5	43.4	52.0	4.2	13.0	20.0
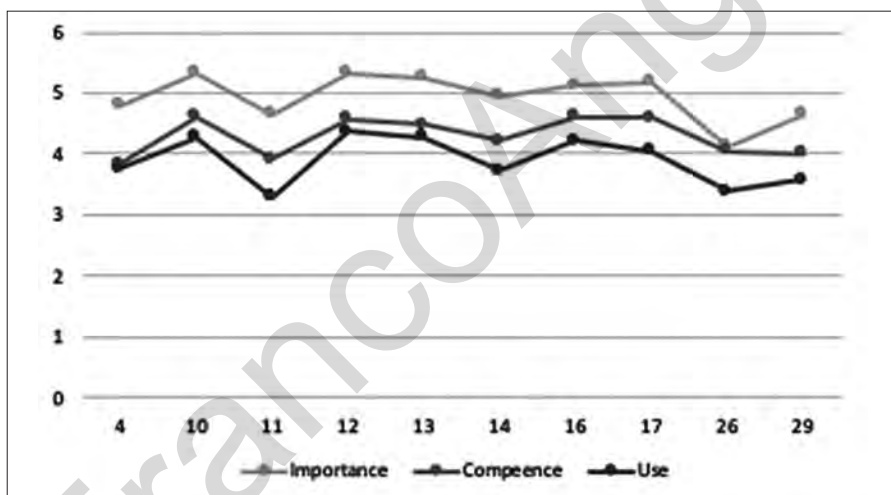
Figure 7 – *Mean per criterion for each item of the promotion of student participation in assessment dimension in Catalonia*

4.2.4. Dimension 4: reviewing, improving and innovating assessment

In the results for *reviewing, improving and innovating assessment* dimen-sion, update knowledge on learning assessment (7) scores high in *impor-tance*, but medium or even low in *competence* and *use*. This could explain the medium-low score on the *competence* and *use* criteria of introducing improvements in assessment (5, 6). As much as participants perceive them-

selves as *competent* and *use* critically analyze the information derived from the assessment processes (24), it seems to be difficult to put this into practice. At the same time, co-ordination between teachers (30) often takes place and they perceive themselves as *competent*.

Table 9 – *Frequency as a percentage of the importance, competence and use criterion for each item of the reviewing, improving and innovating assessment dimension in Catalonia*

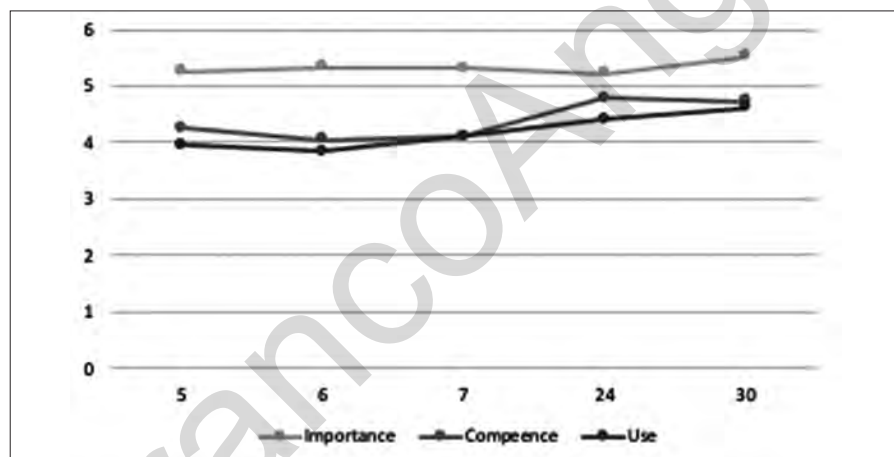| Item	Very often-completely			Sometimes-quite often			Never-very little		
	I	C	U	I	C	U	I	C	U
5	73.5	30.4	48.0	12.5	56.5	52.0	–	–	4.0
6	83.3	34.8	16.0	16.7	47.8	68.0	–	17.4	12.0
7	75.0	34.4	36.0	12.5	43.4	28.0	4.2	8.7	16.0
24	70.8	47.8	36.0	8.3	30.4	36.0	–	4.3	8.0
30	83.4	56.5	56.0	12.5	30.4	28.0	–	8.7	12.0

Figure 8 – *Mean per criterion for each item of the reviewing, improving and innovating assessment dimension in Catalonia*

4.3. Differences between regions

4.3.1. Dimension 1: assessment design and planning

In general, in the *assessment design and planning* dimension, the Catalan teachers perceive a low competence that is manifested in *use*, whereas the Tuscan teachers show greater stability in the three criteria of *importance*, *competence* and *use*.

47

We observed a statistically significant difference between the two regions on items 8 and 19, which occurred in the three criteria, with a lower balance in Catalonia than in Tuscany, which has a more regular distribution.

The results of items 9 and 20 also show a statistically significant difference. The Catalan teachers attach more importance to them than the Tuscan teachers.

Table 10 – *Student T-test per criterion for each item of the assessment design and planning dimension*

Item	Criterion		
	Importance	*Competence*	*Use*
1. Design assessment systems and procedures (determine what, how and when it will be assessed, as well as other specifications necessary to carry out the assessment: criteria, actions, instruments, etc.)	0.06	0.01	0.24
8. Use assessment procedures and techniques consistent with the different methods and modalities of university education	0.02	0.00	0.00
9. Make the objectives, standards and criteria of education known	0.02	0.38	0.36
15. Relate the assessment system with the objectives of the subject and adapt it to them	0.10	0.16	0.71
18. Build assessment instruments	0.12	0.09	0.13
19. Adapt the assessment to learning contexts in which electronic means are used (blended learning, online, e-learning)	0.00	0.00	0.00
20. Consider assessment in an integrated way in the teaching-learning process	0.03	0.31	0.75
25. Make the qualification procedure known	0.94	0.52	0.36
31. Use technological platforms and tools in the assessment process (such as Moodle, LAMS, Blackboard, etc.)	0.12	0.25	0.97

4.3.2. Dimension 2: monitoring student learning

In *monitoring of student learning* dimension, the results of the two regions are very similar, with a difference in perceived less *competence* and *use* of feedback and feedforward of the Catalan teachers.

A first reading reveals a clear similarity between the two regions, although it seems important to highlight the higher score of Catalan teachers in the *importance* criteria in this dimension.

It is also worth noting the differences between the two regions with regards to items 21, 27 and 28, since although the teachers of both regions see them as very *important*, in Catalonia the percentage is higher.

Table 11 – *Student T-test per criterion for each item of the monitoring of student learning dimension*

Item	Criterion		
	Importance	*Competence*	*Use*
2. Provide students with information about the results of the assessment in such a way that they can reflect on their level of achievement (feedback)	0.16	0.03	0.04
3. Provide students with information about the results of the assessment in such a way that they can modify and improve their performances (feedforward)	0.66	0.03	0.00
21. Carry out initial assessment	0.00	0.15	0.42
22. Carry out continuous assessment (by monitoring the learning level of the students)	0.08	0.64	0.56
23. Carry out final assessment	0.50	0.78	0.71
27. Use assessment as a means to know the students' learning needs and thus be able to respond to them	0.00	0.24	0.91
28. Use assessment tasks (essays, reports, portfolios, etc.) to promote learning	0.03	0.84	0.60

4.3.3. Dimension 3: promoting student participation in assessment

In the *promotion of student participation in assessment* dimension we can note that it is a dimension in which teachers in both regions show less *use*, that is, it is promoted less than the other dimensions.

Certain differences between the two regions in the criterion of *importance* should be noted. The Catalan teachers considered items 13 and 14 to be of greater *importance* than the Tuscan teachers. However, the low *use* of co-assessment in Catalonia is surprising.

There are differences in the criterion of *competence* in items 4 and 11. In Tuscany, their *importance* is greater, and so is their perceived *competence* and their use than in Catalonia, where are lower.

It should be noted that use is much lower in item 26 in Catalonia, than in Tuscany. The same is true for item 29, although the difference in *use* is not very high.

Table 12 – *Student T-test per criterion for each item of the promotion of student participation in assessment dimension*

Item	Criterion		
	Importance	*Competence*	*Use*
4. Teach students how to assess and train them in it	0.23	0.00	0.00
10. Make students aware of the benefits of their participation in the assessment practice	0.19	0.96	0.43
11. Encourage student participation in the design of the assessment	0.43	0.02	0.00
12. Encourage student participation through self-assessment (student or group assessment of their activities and performances)	0.31	0.47	0.49
13. Encourage student participation through peer assessment (assessment by students or groups of the activities and performances of their peers)	0.02	0.62	0.32
14. Encourage student participation through co-assessment (teacher and student assess in a consensual and negotiated way the performance or tasks of the student)	0.02	0.44	0.85
16. Agree or reach a consensus with students on what is to be assessed (determine what is going to be assessed: oral communication, autonomous learning, knowledge of basic concepts, etc.)	0.51	0.72	0.20
17. Agree or reach a consensus with students on the assessment criteria (clarity of exposition, relevance and adequacy of the activities carried out autonomously, terminological precision, etc.)	0.35	0.76	0.10
26. Agree or reach a consensus with students on the grading procedure	0.01	0.05	0.00
29. Give students examples and good practices of the assessment tasks carried out by other students, or provide model examples	0.78	0.41	0.12

4.3.4. Dimension 4: reviewing. improving and innovating assessment

In the results for the *reviewing, improving and innovating assessment* dimension, it should be noted that both regions consider item 30 to be of great importance in both regions. However, if we compare the two regions, collaboration and coordination among teachers is promoted more in Catalonia than in Tuscany.

As in the other dimensions, the perceived *competence* of the teachers and the *use* of the different items is closer to the *importance* given to it in Tuscany than in Catalonia, where the item in which it comes closest is the 24.

Table 13 – *Student T-test per criterion for each item of the reviewing, improving and innovating assessment dimension*

Item	Criterion		
	Importance	*Competence*	*Use*
5. Introduce improvements in the assessment processes based on their monitoring	0.25	0.28	0.06
6. Introduce innovations in assessment activities	0.08	2.18	0.05
7. Update knowledge on learning assessment	0.59	0.57	0.21
24. Critically analyze the information derived from the assessment processes	0.38	0.63	0.37
30. Encourage coordination and collaboration among teachers in the assessment processes	0.02	0.54	0.28

In general, the dimension with the greatest differences within the three criteria is the first (*assessment design and planning*), followed by the third (*promotion of student participation in the assessment practice*), the second (*monitoring of student learning*), and the fourth (*reviewing, improving and innovating assessment*) in both regions.

5. Conclusions

In the comparison of the two regions, a considerable difference between Catalonia and Tuscany can be observed, especially since the three criteria are more aligned in Tuscany than in Catalonia. One reason that could be pointed out is the level of governmental prescription in Italy, which can make practices more standardized and bureaucratized than in Catalonia, where it is the schools themselves that develop their own assessment proposals. Hence the need for alignment between policy guidelines and the assessment practices of the schools in order to achieve coherence and continuity within the teaching-learning practices (Caliskan & Zhu, 2019; Darmody, Lysaght, & O'Leary, 2020).

At the same time, for a greater alignment between the whole teaching-learning-assessment process (Cizek, Kosh, & Toutkoushian, 2018), and to prevent the lack of perceived *competence*, demotivation and stress (Flores 2012; Jiang, Sporte, & Luppescu, 2015), while ensuring proper *use* of planning and deployment of the teaching assessment practices, changes in the relationship between teachers and between teachers and the management team could be introduced (Kraft & Gilmour, 2016).

With regards to teachers, the implementation of a co-construction and co-operation of the planning and deployment of assessment practices, especially

51

in Tuscany, would make them more aware of new ways of accomplishing this and/or help consolidate what they are already doing (Folch, Córdoba, & Ribalta, 2020; King-Sears et al., 2020).

With regards to the relationship between teachers and management, the latter should pay attention to the needs of the former, create spaces for sharing, collaborating and cooperating (Folch, Córdoba, & Ribalta, 2020; OFSTED, 2008; Solheim, Roland, & Ertesvåg, 2018) and provide an updated training on assessment practices to improve them (Caliskan & Zhu, 2019; Quesada, Rodríguez, & Ibarra, 2017).

Introducing these changes could lower the stress levels of teachers and thus increase their self-confidence, their perception of competence and use, and their professional and personal development (Solheim, Roland, & Ertesvåg, 2018).

At the same time, as Caliskan & Zhu (2019) point out, the needs of the students must be listened to and their participation in the planning and implementation of the assessment practices increased by various means: a) promoting their co-participation, b) using the three types of assessment equitably, c) diversifying the structures of participation among teachers, d) providing more accurate feedback and feedforward.

The introduction of these elements promotes individual and team metacognition in the students and increases their performance, their self-organization and self-responsibility in their current and future learning process (Adachi, Hong-Meng Tai, & Dawson, 2017; Analí, Paoloni, & Donolo, 2017; Cornoldi, De Beni, & Gruppo, 2020; Li & Grion, 2019; Lucisano & Stanzione, 2018; Murillo & Duk, 2015; Trinchero, 2017). It also increases their perception of the teaching-learning-assessment process as a whole and thus introduces new regulatory strategies (Koenen, Dochy, & Berghmans, 2015).

One limitation of this study is the difference in the number of participants in the two regions, being much higher in Tuscany than in Catalonia. It is also important to point out, as already mentioned, that there is no instrument like ActEval (Quesada, Rodríguez, & Ibarra, 2013) that meets the three criteria of assessment practices at educational levels before university, although this instrument can be also applied to these levels.

Future studies could compare other regions of other countries or the same country, in order to broaden the perspective on assessment practices and further investigate the needs of teachers regarding this process.

References

Adachi, C,. Hong-Meng Tai, J. & Dawson, P. (2017), "Academics' perceptions of the benefits and challenges of self and peer assessment in higher education", *Assessment and Evaluation in Higher Education*, 43 (2), pp. 294-306, doi: 10.1080/02602938.2017.1339775.

Analí, V., Paoloni, P.V., Donolo, D.S. (2017), "Creencias de autoeficacia y contextos de evaluación. Un estudio con estudiantes universitarios", *Revista actualidades investigativas en educación*, 17 (2), pp. 112-143. doi: 10.15517/aie.v17i1.28144.

Black, P.J., William, D. (2018), "Classroom assessment and pedagogy", *Assessment in Education: Principles, Policy & Practice*, 25 (3), pp. 551-575, doi: 10.1080/0969594X.2018.1441807.

Caliskan, A., Zhu, C. (2019), "Organizational culture and educational innovations in Turkish high education: perceptions and reactions of students", *Educational Sciences: Theory & Practice*, 20 (1), pp. 20-39, doi: 10.12738/jestp.2020.1.003.

Ciani, A., Ferrari, L., Vannini, I. (2020), *Progettare e valutare per l'equità e la qualità nella didattica. Aspetti teorici e indicazioni metodologiche*. Milano: FrancoAngeli.

Cizek, G.J., Kosh, A.E., Toutkoushian, E.K. (2018), "Gathering and evaluating validity evidence: the generalized assessment alignment tool", *Journal of Educational Measurement*, 55 (4), pp. 477-512, doi: 10.1111/jedm.12189.

Cornoldi, C., De Beni, R., Gruppo, M.T. (2020), *Imparare a studiare: strategie, stili cognitivi, metacognizione e atteggiamenti nello studio*. Trento: Centro Studi Erickson.

Darmody, M., Lysaght, Z., O'Leary, M. (2020), "Irish post-primary teachers' conceptions of assessment at a time of curriculum and assessment reform", *Assessment in Education: Principles, Polixcy & Pratice*, 27 (5), pp. 501-521, doi: 10.1080/0969594X.2020.1761290.

Flores, M.A. (2012), "The implementation of a new policy on teacher appraisal in Portugal: how do teachers experience it at school?", *Educational Assessment Evaluation and Accountability*, 24 (4), pp. 351-368, doi: 10.1007/s11092-012-9153-7.

Folch, C., Córdoba, T., Ribalta, D. (2020), "La performance: Una propuesta interdisciplinar de las áreas de educación física, educación musical y educación visual y plástica en la formación inicial de los futuros maestros", *Retos*, 37, pp. 613-619, doi: 10.47197/retos.v37i37.74187.

Jiang, J.Y., Sporte, S.E., Luppescu, S. (2015), "Teacher perspectives on evaluation reform", *Educational Researcher*, 44 (2), pp. 105-116, doi: 10.3102/0013189 X15575517.

King-Sears, M.E., Strogilos, V. (2020), "An exploratory study of self-efficacy, school belongingness, and co-teaching perspectives from middle school students and teachers in a mathematics co-taught classroom", *International Journal of Inclusive Education*, 24 (2), pp. 162-180, doi: 10.1080/13603116.2018.1453553.

Koenen, A., Dochy, F., Berghmans, I. (2015), "A phenomenographic analysis of the implemetnation of competence-based education in higher education", *Teaching and Teacher Education*, 50, pp. 1-12, doi: 10.1016/j.tate.2015.04.001.

Kraft, M.A., Gilmour, A.F. (2016), "Can principals promote teacher development as evaluators? A case study of principals' views and experiences", *Educational Administration Quarterly*, 52 (5), pp. 711-753, doi: 10.1177/0013161X16653445.

Kulasegaram, K., Rangachari, P.K. (2018), "Beyond 'formative': assessments to enrich student learning", *Advanced in Phycology Education*, 42, pp. 5-14, doi: 10.1152/advan.00122.2017.

Li, L., Grion, V. (2019), "The power of giving feedback and receiving feedback in peer assessment", *All Ireland Journal of Higher Education*, 11 (2), pp. 1-17.

Lucisano, P., Stanzione, I. (2018), "Ripensare il successo scolastico a partire dal vissuto degli studenti", *Ricercazione*, 10 (2), pp. 79-100, doi: 10.32076/RA10206.

Murillo, F.J., Duk, C. (2015), "Escuelas inclusivas: gestión para el cambio y la mejora", *Diplomado en inclusión educativa escuelas inclusivas: enseñar y aprender en la diversidad*, OEI-CAEU.

OFSTED (2008), *Assessment for learning: the impact of National Strategy support*. OFSTED.

Quesada, V., Rodríguez, G., Ibarra, M.S. (2013), "ActEval: un instrumento para el análisis y reflexión sobre la actividad evaluadora del profesorado universitario", *Revista de educación*, 362, pp. 69-104, doi: 10.4438/1988-592X-RE-2011-362-153.

Quesada, V., Rodríguez, G., Ibarra, M.S. (2017), "Planificación e innovación de la evaluación en educación superior: la perspectiva del profesorado", *Revista de investigación educativa*, 35 (1), pp. 53-69, doi: 10.6018/rie.35.1.239261.

Solheim, K., Roland, P., Ertesvåg, S.K. (2018), "Teachers'perceptions of their collective and individual learning regarding classroom interaction", *Educational Research*, 60 (4), pp. 459-477, doi: 10.1080/00131881.2018.1533790.

Trinchero, R. (2017), "Nove concetti chiave per un'istruzione informata dall'evidenza", *Formazione & Insegnamento*, 15 (2), pp. 113-125, doi: 107346/-fei-XV-02-17_11.

Vannini, I. (2011), Dalla diagnosi in ingresso alla valutazione sommativa. Il controllo degli apprendimenti per promuovere la cultura della valutazione all'interno delle scuole. In D. Capperucci (a cura di), *La valutazione degli apprendimenti in ambito scolastico. Promuovere il successo formativo a partire dalla valutazione*. Milano: FrancoAngeli, pp. 40-62.