Chapter 5
The Use of a Rating Scale as a Formative and Shared Assessment Tool in Physical Education

Daniel Bores-García

https://orcid.org/0000-0003-2522-8493

Rey Juan Carlos University, Spain

Raúl A. Barba-Martín

University of León, Spain

Gustavo González-Calvo

University of Valladolid, Spain

David Hortigüela-Alcalá

https://orcid.org/0000-0001-5951-758X

University of Burgos, Spain

ABSTRACT

This chapter presents the analysis of a formative and shared assessment experience in the subject of physical education in secondary education in a high school in Spain. In the assessment process, an assessment scale has been used as an assessment instrument, by means of which students have self-assessed themselves, have been co-assessed by their peers and have received a hetero- assessment from the teacher based on the criteria previously established in the instrument. After the implementation of this experience, a study was carried out on the students' perception of the formative and shared assessment process and the use of the evaluation scale. A discussion group was held, and the teachers' and students' diaries were analyzed. The results show a feeling of motivation and commitment in the students to the task and the group thanks to the feeling of being part of the process through the assessment.

ASSESSMENT, GRADING, AND BODY EXPRESSION

The relationship between assessment and body language as content in the subject of Physical Education (from now on PE) has been a problem for teachers in recent years (Bores-García et al., 2021). On the one hand, the content itself is conflicting from a pedagogical point of view. A careful analysis of this tension reveals three frequent errors that cause this conflict (Lafuente & Hortigüela-Alcalá, 2021): firstly, not developing content that is associated with the collective achievement of the class. Secondly, not generating satisfaction in one's own bodily experiences. Finally, not implementing coherent proposals of a longitudinal nature. This last aspect is one of the main limitations of the projection of these contents since, as they are not applied with a certain logic and continuity, they do not allow learning outcomes associated with the interventions carried out to be obtained (Mattsson & Lundvall, 2015). In this way, it becomes especially complex to generate positive experiences in students since if the PE teacher himself does not give relevance to these contents, students cannot be expected to be motivated towards them (MacLean, 2018). On the other hand, and associated to the difficulty to work on these contents, is knowing how to evaluate them. One of the main problems that exist among PE teachers is the constant identification and homologation of assessment with grading (Álvarez, 2005; Hortigüela-Alcalá et al., 2019a). It should be possible to assess without having to grade everything, and this is a fact that generates many problems in the educational community, especially when working on the contents of PE, which cannot be based solely on surpassing pre-established marks. Assessment, especially in PE, must be synonymous with learning and, for this, the planning must be constructed from the beginning together with the methodological, organizational aspects and the tasks set (Brevik et al., 2017). However, the problem is that when this process is not carried out coherently with the contents to be worked on, the assessment and grading process tends to be hidden from the students (Barba-Martín & Hortigüela-Alcalá., 2022. . In this sense, another problem associated with the confusion between assessment and grading lies in the use of assessment as a power mechanism of the teacher, who justifies being the only one competent to assess because he/she is the one with the knowledge (Fernández-Balboa, 2005, 2007). When this happens, the essence of teaching is lost, as students are not allowed to be an active part of a process of which they are a substantial part and it is difficult for them to develop important aspects of their learning such as autonomy, self-regulation, motivation and awareness of what they have learned((Leenknecht et al., 2020; Ozan & Kincal, 2018). This means that the body does not acquire all the necessary pedagogical dimension within the subject, and that on many occasions the result only responds to physiological criteria and is based strictly on motor performance (López-Pastor et al., 2013). A subject and assessment approach That is detrimental to students and their subsequent physical development outside the classroom (Beltrán-Carillo & Devís-Devís, 2019).

In view of this, a formative and shared approach to assessment are ideal for promotes aspects that relate to and encourage learning such as self-regulation, motivation, responsibility and awareness of learning on the part of the student on the one hand (Hortigüela-Alcalá et al., 2015: Leenknecht et al., 2020; Weldmeskel et al., 2016), and the extrapolation of learning to different contexts on the other (Chng & Lund, 2018). In this line, in the Spanish context in which this work is framed, in the spanish context, in which this work is framed several studies in recent years have shown the benefits associated with these aspects of formative and shared assessment in primary and secondary education (Bores-García et al., 2020; Heras-Bernardino & Herrán-Álvarez, 2019; López-Pastor & Pérez-Pueyo, 2017; Bores-García et al., 2021) and in higher education (Barba-Martín et al., 2020b; Hamodi et al., 2018; Hamodi & Barba-Martín, 2021; Hortigüela-Alcalá et al., 2019b; Romero et al., 2017)., In this way, it is

proven how fundamental it is that in the development of a PE centered on working on the multitude of possible learning of this subject, such as corporal expression, the PE teachers should use assessment as a tool with a clear methodological component, focused on learning awareness and not only on the final grade (Hortigüela-Alcalá et al., 2021). To this end, a clear assessment system and instruments must be developed from the outset (Pegalajar, 2021).

In this sense, the present work tries to contribute a new experience to the literature on assessment in Physical Education and, for this reason, the main objective of this chapter is the narration of a school experience in which formative and shared assessnet was used through the use of an rating scale in a Didactic Unit (DU) of body expression

FORMATIVE AND SHARED ASSESSMENT: AN ESSENTIAL PART OF THE TEACHING-LEARNING PROCESS

But what is formative assessment? and shared assessment? How is it carried out in the teaching-learning process in PE? Formative assessment is an approach focused on the use of this aspect, within the teaching-learning process, to help students achieve quality learning through feedback. (Biggs, 2005; Brown & Glasner, 2003). This assessment model stands out because it is focused on achieving a conscious construction of learning on the part of the students, rather than only on the final result. (López-Pastor et al., 2013). In this sense, it is important to understand that there is no point in grading learning at the end of the process if students has not given the opportunity to learn from their mistakes and to reinforce their successes in order to progress (Brown & Pickford, 2013; Santos Guerra, 2014), which is still too often the case today (Panadero et al., 2019).

On the other hand, shared assessment is a model of assessment that, by understanding this as a formative process, understands as necessary and important the active participation of the students (Barba-Martín & Hortigüela-Alcalá, 2022), based on dialogic, ethical, democratic and transparent relations between the different agents (Santos Guerra, 2003). The learner must be aware of his or her learning and must therefore be involved in the assessment process through different participatory mechanisms such as self-assessment, as the process through which the students or the group is able to evaluate their work at a given moment (Boud, 2013), or peer assessment, where peers give feedback to their peers (Topping, 2009). The possibility for all actors to participate in the assessment of learning processes provides the opportunity for learners to receive feedback from different points of view that help them to become aware of their learning and to be able to direct it. However, it is necessary to establish clear assessment criteria and to create critical scaffolding processes with this type of assessment in the student body, so that all feedback is relevant (Rust, Price & Donovan, 2003). . Under this perspective, the whole assessment process and the instruments must be built, the latter must be clear and provide the students with relevant information from the beginning (Fraile et al., 2017). Some of the most commonly used instruments are rating scales, since they allow students a progressive approach to learning under the construction of evaluation criteria sequenced in different levels of achievement (López-Pastor & Pérez-Pueyo, 2017). Thus, students can know at any moment of the learning process where they are and make decisions based on it. In the field of Physical Education, these evaluation criteria and their levels of achievement must include all areas of student development (motor, cognitive and social) in a related manner. This has provoked a debate in the subject, under the perspective that the importance of the motor aspect is lost. However, this is a false belief supported by the tradition of grading motor contents such as physical condition through tests

and which should be broken in order to provide students with learning processes based on the evaluation and construction of complete motor experiences (Barba-Martín et al., 2020a).

The instruments should be given to the students at the beginning of the requested activity, and their understanding of the criteria and the established achievements should be contrasted with them (Fraile et al., 2017; Pegalajar, 2021). In this way, students can work with the instrument from the beginning and make conscious decisions during the teaching-learning process through self-assessment processes (Nicol & Macfarlane-Dick, 2007). Regarding the feedback received through peer or teacher assessment, the comments received by the students should always be focused on learning, based on the evaluation criteria of the instrument, and with a clear pedagogical intention (Sadler, 2010). Therefore, as can be seen, the use of evaluation instruments with a formative and shared approach is beneficial for student learning; however, despite its advantages, there are still not many experiences and a traditional style of assessment continues to predominate.

A PRACTICAL EXPERIENCE IN PE AT THE SECONDARY EDUCATION STAGE

The work has been developed with 253 students of 1stESO (12-13 years old) in the 2018/2019 academic year and with 182 students of the same grade in the 2020/2021 academic year. The rating scale, inspired by the proposal of Hernando-Garijo, Hortigüela-Alcalá and Pérez-Pueyo (2017), is used during the process of creating and rehearsing a final group performance based on corporal expression and as an instrument of formative evaluation and grading of the DU, using this instrument to perform antriadic assessment (self-assessment, co-assessment and hetero-assessment). The objective of this triangulation of assessments during the DU is for students to contrast the different perceptions of their work in order to regulate their learning in view of the final production. It was decided to use this rating scale because of its high didactic and methodological component. This scale has been extracted from the website of the Attitudes Group (https://grupoactitudes.com/documentos/), which is widely consulted by teachers. This rating scale has been used in many educational contexts, always under the premise of generating awareness and learning in PE students. It should be remembered that the main objective of the research is to assess the experience of PE students when using this assessment instrument, focusing on a triple feedback: self-assessment, co-assessment and heteroassessment. In this case, it has been used in the content of corporal expression. Previous experience in this sense (Pérez-Pueyo et al., 2017;2019) has shown the positive effects of triadic assessment on student involvement and self-regulation of tasks.

The experience narrated is part of the Unit "I use my body to tell stories", belonging to the 1st ESO programme of the subject of PE and timed in the second assessment of the course, between the months of January and February of both years. This DU has been applied in all the groups of 1st ESO of the IES Alameda de Osuna, in the district of Barajas (Madrid, Spain). The unit lasted four weeks, with two 55-minute lessons per week. For its explanation we have divided the DU into phases, according to the assessment work carried out.

Phase 1. Construction of experiences with corporal expression. During the first three sessions, the teacher proposed a series of imitation, representation and expression activities, in which the students worked first individually and then in small groups. The main aim of these first introductory sessions was to arouse the pupils' curiosity about the expressive capacities of the body and to encourage them to explore all the possibilities that could be achieved through movement. Moreover, being the first year of a new educational stage, there was a great deal of heterogeneity in the pupils from the point of view of

their school of origin. This meant that some pupils had deep and positive experiences of content related to body language, while other pupils were confronted almost for the first time with content that not everyone accepts in the same way.

Phase 2. Delivery and explanation of the assessment criteria through the instrument. In lesson 4 the teacher explained to the students the process they were about to start. He detailed the phases of the process and handed out the assessment scale (Table 1), explaining each of the assessable items and answering any doubts that arose in each class. After the explanation, the students were allowed to create their own working groups following two essential criteria: the groups had to be mixed, including at least two boys and two girls, and they could not exceed a maximum number of participants so that there would be five working groups in each class. In the remaining minutes the pupils were allowed to start discussing the first ideas of the creative process.

Phase 3. Creation of the production based on the criteria of the instrument and teacher feedback. During lessons 5, 6 and 7 the pupils created and rehearsed their final performances in groups. Each group had to bring the rating scale to each session, so that they could read the aspects to be evaluated and thus the most important elements to be taken into account.

Table 1. Rating scale

Name and surname of the group members:				
Prerequisites: groups of 4-5 students (mixed). All participants will be involved in the creative process, rehearsals and performance. Three lessons will be dedicated to the creation and rehearsal of the performance.				
Valuable items	**Subvalue**	**Achievement criteria**		**Maximum value**
Storyboard creation	15	The storyboard has good drawings and collects the most important elements of the representation.		15
	10	The storyboard has good drawings and includes some elements of the representation.		
	10	The storyboard has little worked drawings and collects the most important elements of the representation.		
	5	The storyboard has little worked out drawings and picks up some elements of the representation.		
	0	The storyboard has very poor drawings and provides almost no information.		
Duration	10	The performance lasts between 4 and 6 minutes		10
	5	The representation deviates less than 30 seconds from the stipulated time.		
	0	The representation deviates more than 20 seconds from the stipulated time.		
Narrative capacity	20	The story is perfectly understandable		20
	12	The story is understood, although not completely		
	5	There are many elements that are not understood, which makes it difficult to follow the story		
	0	The story is incomprehensible		
Expressive capacity	20	The representation is very rich in expressive body elements		20
	12	The representation has some expressive body elements		
	5	Representation has few interesting expressive resources		
	0	The representation does not have any interesting expressive resources		
Creative capacity	20	The performance is very original and creative		20
	12	The performance has some interesting creative aspects		
	5	The representation has almost no creative elements		
	0	The performance is neither original nor creative		
Flow (coordination)	15	There is coordination and rapport between the members of the group		15
	8	Coordination and rapport can be improved		
	0	The members of the group act in an uncoordinated way, without empathy.		
FINAL SCORE (maximum 100 points)				

Each group worked in an autonomous way and the teacher went around the groups solving doubts, contributing ideas and encouraging those groups that were a bit confused in the creative process.

Phase 4. Peer and heteroassessment process with the instrument. In lesson 8 the students performed the performance and the classmates and the teacher had the opportunity to give feedback based on the

rating scale. At the end of each group's performance, the teacher gave a rating scale to each group to evaluate the group they had performed. He also gave the group a document to evaluate their own performance. Finally, the teacher evaluated the group as well. A few minutes were dedicated to generate a process of student participation with the objective of letting each group know those elements in which they could improve, according to what was expressed in the assessment scale.

Phase 5. Reconstruction of the productions based on the feedbacks received. Lesson 9 was aimed at correcting those elements that could be improved, in order to improve the representation for the following lesson.

Phase 6. New assessment process with the instrument and jump to qualification. Finally, in lesson 10, the final presentation of the performances took place, after each of which the same process took place as in session 8. At the same time, the teacher evaluated each group's performance. A few minutes were devoted to a discussion of the elements assessed, after which the qualitative assessments (achievement criteria) were transferred to the score assigned to each of them, making the final sum out of a total of 100 possible points. In this way, the leap from assessment to marking was made, obtaining a final average mark after equal weighting of the mark obtained in the self-assessment, the co-assessment and the teacher's assessment.

In addition, students completed a semi-structured personal learning diary during some days of the unit, in which they answered some questions proposed by the teacher for each session. Day 1 of the personal diary was the last rehearsal before the first performance, day 2 was the day of the performance, which the pupils themselves called "mock" as it was not the final performance, day 3 was the day of the final performance and day 4 was completed at the end of the didactic unit. The following table shows some of the questions that the students had to answer in each session of the didactic unit.

Table 2. Student´s personal diary questions

Suggested questions for each lesson to complete the diary
Day 1. What motivates you most to do a good performance? Knowing that your classmates and the teacher are going to evaluate you, what does it feel like?
Day 2. Has doing this performance simulation helped you? What things have helped you? What things are you going to change and improve?
Day 3. Describe what you think your level of commitment and work in the group has been during the process of creation and rehearsal. Also give your opinion about the level of commitment and work of your colleagues. What have you learned during the process of creation and rehearsal? How has it been useful for you to have the rating scale during the whole process?
Day 4. How have you experienced having to evaluate your classmates? Did you find it difficult? Did you try to benefit or harm them, or did you remain as objective as possible? How have you experienced having to evaluate your own performance? Did you find it difficult? Do you think the grade given to you by your classmates was fair? Why or why not? Do you think the grade given to you by the teacher was fair? Why or why not?

The PE teacher who carried out the teaching of the didactic unit completed a personal diary with reflections on what happened in each of the sessions with the different groups. Finally, two focus groups were conducted with eight of the participating students.

Table 3. Basic script for the discussion groups

Basic script for the discussion groups
How did you live knowing that you were going to be evaluated in three different ways?
What has been your motivation for trying to do well?
Has the mock performance been helpful in improving performance? In what?
Do you think there have been differences in commitment among the members of your group?
How have you been able to evaluate your peers, evaluate yourself and know that you were being evaluated by the teacher?
Do you consider the score obtained by each of the three ways to be fair?

In order to facilitate the reading of the conclusions obtained from the experience and from the personal diaries and the discussion groups, we will proceed to present the most relevant elements exemplified with some extracts from the teacher´s diary (TD), the student´s personal diaries (SPD) and the focus groups (FG). A qualitative methodology was used to gain in-depth knowledge of the students' evaluations of the use of the scale and their involvement in its evaluation. To this end, the research was carried out in the context of the classroom, so as not to distort the real day-to-day practice of the students, thus connecting with their most personal and social dimension (Halquist & Musanti, 2010). Data were triangulated between the three data collection instruments used. This made it possible, through the use of cross-patterns, to assign each text extract to a category of results, thus obtaining the saturation of ideas (Tortorella et al., 2015). The allocation of data to each of the categories and subcategories was thoroughly reviewed by each of the researchers independently. This process allowed for a coding of the data with each of the instruments used, thus showing linearity and coherence in the analysis (Del Río-Roberts, 2011).

1.   Increased motivation in the use of formative and shared assessment

One of the most remarkable elements was motivation. First of all, they were very motivated by the possibility of forming their own groups, even if there were some requirements they had to respect.

One of the things that made them most excited when I explained the working procedure was when I told them that they would have a few minutes to form their own groups, as long as they respected the grouping criteria (TD).

Obtaining a good mark that would have an impact on their final grade in the course was also a motivating factor. The prospect of having part of the responsibility for this grade increased their motivation, as it is usually the teacher who is responsible not only for the assessment but also for the marking of the tests given during the course in the different subjects.

The mark also motivates me, because I want to get a good mark like in the first evaluation (SPD).

Knowing that two thirds of the grade came from what we put on ourselves and the grade our classmates put on us, that was quite motivating (FG).

Finally, the pupils were attracted by the content. Although in each class there is always a group of pupils who relate PE exclusively to sport, they were generally enthusiastic about being able to do a

teaching unit on body language. Even more so as the two previous teaching units had been on handball and physical fitness.

I like being able to show that physical education is more than just sport (SPD).

I was a bit fed up with handball and athletics. I love theatre and that makes me want to do my best for the first time in the course (SPD).

2.   Group feedback and second chances

Students are not used to being able to take an assessment test without consequences for the grade. However, formative assessment is based on the idea of not always having to link assessment with marking. Therefore, students were given the opportunity to evaluate themselves and their peers without the pressure of marking.

Being able to do a performance test before the final was very good for us to know what things we needed to improve and make it perfect in the final assembly (FG).

The students acknowledged that they found this session useful, as they identified shortcomings in their expressive creations that they were able to remedy in the following session.

We realized that the performance was too short and did not meet the minimum time requirement. It's one thing when you do it in rehearsal and another when you do it in front of an audience. When you're nervous, you go faster than planned (SPD).

This possibility allowed them to face the performance of the final session with greater confidence and security, knowing that with the corrections and improvements made, the result would always be better than that of the mock session. In addition, the students who had difficulty in being in front of an observing audience were able to have a first experience in order to face the second one with greater confidence.

Because you've already performed in the simulation, you feel safer and more confident to do it in front of everyone (SPD).

It's helped me to get used to performing in public, because it's not the same to do it only when everyone is looking at you (SPD).

The assessment scale has been an advantage for the pupils, who from the beginning of the project to create the representation were able to know which aspects could be assessed and, therefore, which elements they had to focus on in order to create the representation correctly. The pupils themselves admit to having used the rating scale to detect errors and how to solve them.

The rating scale has been very useful to see where the failures were and what we could do to correct them (FG).

However, the teacher acknowledges that some students did not make correct use of the rating scale, although it has been advantageous for those who have used it correctly.

Some of the students had not read the rating scale carefully, even though we spent almost a whole lesson reading and explaining it. Having to use it to assess themselves and their peers has been good for the next two sessions (TD).

3.   Group commitment

As with any group project, problems arise during its development, the resolution of which is an essential part of the project itself and of interdisciplinary and transdisciplinary learning of great educational value. Many of the groups took the expressive project very seriously, to the point that some of them created Whatsapp groups exclusively to talk about the Project.

...we even have a Whatsapp group to talk to and discuss the failures... (SPD).

The students admit that having common objectives has made them aware of the need to contribute individually to the group. Knowing that assessment and marking were carried out as a group has awakened the need to work as a team, knowing that the consequences of individual actions would have repercussions for the group.

The overall level of involvement has been high. Having shared elements such as a common project, a common evaluation and a common rating has served as a glue for individual and group commitment (TD).

All my classmates have also committed themselves and worked just as hard as I did (SPD).

However, other students have complained about the low level of group commitment shown by some members of the group, who have not contributed to the same extent and have instead hindered the process of creation and rehearsal.

There are people who have had a hard time taking the piss out of it (SPD).

As is always the case, there is a small percentage of students with little motivation for the task (and for everything in general too, almost always) who either blow up the work of others or remain absent, blowing it up anyway. This has occurred above all in groups of students who have not been able to group with whomever they wish, resulting in a group that is too heterogeneous, with students who are more marginalized in the class and, therefore, less motivated (TD).

4.   Lessons learned during the process

Most of the pupils refer to social or relational learning, as the most intense process from an educational point of view has been group work and the resolution of problems that occurred along the way.

It has helped me to know more about my classmates and that some people are not what they seem, because they can be very funny, intelligent and even very good people (SPD).

During the process of creation and rehearsal I have learned to listen (SPD).

This learning is not accidental, but the teacher himself who carried out the didactic unit recognises the intention behind it.

I myself have expressly decided not to give so much importance to expressive content, but to use it as a means of achieving other elements of equal or greater value such as respect for oneself and others, effort, mutual commitment, honesty and transparency (TD).

Working in a team forces students to listen to the opinions of others and not to impose their own. The development of empathy has been a key part of the process.

During the process of creation and rehearsal I have learned to listen (SPD).

5. Lessons learned during the process

One of the key elements of formative evaluation has been the use of the rating scale, as has already been pointed out several times in this chapter. Its presence throughout the process has allowed the groups to work more efficiently, more focused on the expected results and with greater organisational capacity.

It was quite useful to me because that way we could see everything we had to do and so we were quite focused. We've organized ourselves much better than if we'd had nothing (SPD).

The teacher acknowledges, however, that it is not easy to convey to students the importance of using this assessment tool, especially in this first year of secondary school where some students have already used this type of instrument while for others it is a complete novelty.

Trying to show the students the importance of the rating scale has been difficult for me, I admit. They're not used to using paper and pencil in P.E. Some groups forgot to bring the rating scale to the rehearsals and I got tired of reminding them of its importance. But in the end I think it has been very useful for them (TD)

The lack of previous experience in formative and shared assessment has hindered the process of assimilation of this way of working, because students in the first years of secondary school may misuse shared assessment if they are not used to it.

Evaluating my classmates has been a little difficult, because while I wanted to help them I also wanted to be objective (SPD).

Assessing my own performance was strange to me, because we didn't see what we were doing ourselves (SPD)

However, the rating scale has helped them in this process, especially in the aspect of objectivity or thoroughness in evaluations and feedback.

With the rating scale it has been easier to evaluate my colleagues and myself (SPD).

Following the sheet (rating scale) has helped me stay objective (SPD).

Finally, the leap to marking has also involved some conflicts produced by the inexperience of the students in these assessment processes. The less mature students with greater difficulties of commitment and behaviour did not correctly assess and mark their own work or that of their peers, while the majority of students showed great responsibility and maturity in trying to adjust as much as possible to the assessment scale in order to make judgements and to make the leap to the mark. One of the biggest problems for students has been objectivity. Interestingly, and perhaps only by comparison, students point out that the teacher is objective because he/she has the most knowledge on the subject, whereas they cannot be completely objective.

The teacher's grade did seem fair to me because it was the most objective of all (SPD).

I think I have been fair because the grade we have given the other groups has coincided with the grade those groups have given themselves and with the grade the teacher has given them (SPD).

I recognize that we have scored ourselves a little higher than we really deserved, to help ourselves. The teacher's grade was the fairest (FG).

The concept of fairness appears in the students' diaries and in the discussion groups very often associated with the concept of objectivity, pointing out that the only fair assessment is that of the teacher or any other assessment that resembles the teacher's assessment.

CONCLUSION FROM THE ABOVE EXPERIENCE AND PROPOSALS FOR FUTURE INITIATIVES

The aim of this paper is to narrate an experience of working on corporal expression in Physical Education through the formative and shared use of a graded scale. For this purpose, the process carried out has been narrated in detail, as well as the main reflections of the teachers and students involved during the process. In this way, a new practical experience is contributed to a still brief literature.

After the body expression teaching unit described above, in which formative and shared assessment was used, the results support the findings of other predecessor research, discussed in the first part of this paper. Students reported feeling more motivated and engaged in the task by being part of the assessment process. They recognise the positive sense of knowing from the beginning of the process the elements that would be observed in the final performance, so that during the sessions they can direct their efforts

more effectively and efficiently. In addition, they comment that they have felt more engaged with their group through the use of formative and shared assessment, feeling more involved in the whole process from the first lesson, not only in the final lesson where the leap to marking is made. The use of this type of assessment has allowed them to understand from the very beginning of the process the effect that individual actions have on the rest of the group at an evaluative level develops the group commitment of all members of each team. In addition, a specific contribution of this experience has been with respect to the use of the instrument as a means of scoring by all the agents involved, it showed small differences between their perceptions, the higher levels of achievement were those coming out of the self-assessment processes, followed by the assessment coming out of the teacher assessment. The mechanism that resulted in the lowest levels of achievement was peer assessment, as the students themselves acknowledge that they are more critical in detecting their peers' mistakes than in detecting their own mistakes and try to be more objective, which they believe makes it easier for them to use the instrument.

The use of formative and shared assessment in general and instruments such as rating scales in particular show their educational potential for increasing the motivation and group commitment of PE students, In this case, to content related to bodily expression. the reported experience, as well as previous research, shows that it is clear that it is advisable to involve students in the assessment processes through, among others, instruments applied using triadic assessment. Although this experience highlights how the use of a formative and shared assessment and the use of an assessment scale require prolonged work in the classroom so that students learn to work with objectivity and awareness of the established criteria, both with their own productions and with those of their classmates.

A proposal for future interventions in which formative and shared assessment is used is the joint creation of the assessment instrument, in such a way that the pupils participate even more in the assessment process by having to make decisions on the achievement criteria and this will serve as a scaffolding with the instrument and the criteria, for which they must previously know the content being learnt and priorities some learning over others in order to grade the degree of achievement of the objectives established for the didactic unit. However, it should be borne in mind that this will take at least two or three sessions prior to the start of the creation and rehearsal, so the teacher must assess whether there is enough time for this. In addition, the process will be different for learners with extensive experience in co-creating assessment instruments than for learners with no experience in this area, because as students experience formative and shared evaluation processes on a constant basis and participate actively in the creation of instruments, these times will be reduced thanks to the accumulated experience.

REFERENCES

Álvarez, J. (2005). *Evaluar para conocer, examinar para excluir*. Morata.

Barba-Martín, R. A., Bores-García, D., Hortigüela-Alcalá, D., & González-Calvo, G. (2020). Evaluación formativa con los estudiantes en prácticas para reducir la brecha teoría-práctica en la formación inicial del profesorado. *Educacion Fisica y Deporte*, *39*(1). doi:10.17533/udea.efyd.v39n1a02

Barba-Martín, R. A., Hortigüela-Alcalá, D., & Pérez-Pueyo, A. (2020). Evaluar en educación física: Análisis de las tensiones existentes y justificación del empleo de la evaluación formativa y compartida. *Educacion Fisica y Deporte*, *39*(1). doi:10.17533/udea.efyd.v39n1a03

Barba-Martín, R. A., & Hotigüela-Alcalá, D. (2022). Si la evaluación es aprendizaje, he de formar parte de la misma. Razones que justifican la implicación del alumnado. *Revista Iberoamericana de Evaluación Educativa*, *15*(1), 9–22. doi:10.15366/riee2022.15.1.001

Beltrán-Carrillo, V. J., & Devís-Devís, J. (2019). El pensamiento del alumnado inactivo sobre sus experiencias negativas en educación física: Los discursos del rendimiento, salutismo y masculinidad hegemónica. *RICYDE. Revista Internacional de Ciencias del Deporte*, *55*(15), 20–34. doi:10.5232/ricyde2019.05502

Biggs, J. (2005). *Calidad del aprendizaje universitario*. Narcea.

Bores-García, D., Hortigüela-Alcalá, D., González-Calvo, G., & Barba-Martín, R. A. (2020). Peer Assessment in Physical Education: A Systematic Review of the Last Five Years. *Sustainability (Basel)*, *12*(21), 9233. doi:10.3390u12219233

Bores-García, D., Hortigüela-Alcalá, D., Hernando-Garijo, A., & González-Calvo, G. (2021). Analysis of student motivation towards body expression through the use of formative and share assessment. *Retos*, *40*(40), 198–208. doi:10.47197/retos.v1i40.83025

Boud, D. (2013). *Enhancing learning through self-assessment*. Routledge. doi:10.4324/9781315041520

Brown, S., & Glasner, A. (2003). *Evaluar en la Universidad. Problemas y nuevos enfoques*. Narcea.

Brown, S., & Pickford, R. (2013). *Assessing Skill and Practice*. Routledge.

Chng, L. S., & Lund, J. (2018). Assessment for Learning in Physical Education: The What, Why and How. *Journal of Physical Education, Recreation & Dance*, *89*(8), 29–34. doi:10.1080/07303084.2018.1503119

Del Rio-Roberts, M. (2011). How I Learned to Conduct Focus Groups. *Qualitative Report*, *16*(1), 312–315.

Fernández-Balboa, J. M. (2005). La autoevaluación como práctica promotora de la democracia y la dignidad. En A. Sicilia, & J. M. Fernández-Balboa (Coords.), La otra cara de la educación física: la educación física desde una perspectiva crítica (pp. 127-158). Inde.

Fernández-Balboa, J. M. (2007). Dignity and democracy in the college classroom: The practice of self-evaluation. In R. A. Goldstein (Ed.), *Useful Theory: Making Critical Education Practical* (pp. 105–128). Peter Lang Publishing.

Fraile, J., Pardo, R., & Panadero, E. (2017). ¿Cómo emplear las rúbricas para implementar una verdadera evaluación formativa? *Revista Complutense de Educación*, *28*(4), 1321–1334. doi:10.5209/RCED.51915

Halquist, D., & Musanti, S. I. (2010). Critical incidents and reflection: Turning points that challenge the researcher and create opportunities for knowing. *International Journal of Qualitative Studies in Education : QSE*, *23*(4), 449–461. doi:10.1080/09518398.2010.492811

Hamodi, C., & Barba-Martín, R. A. (2021). *Evaluación formativa y compartida: nuevas propuestas de desarrollo en Educación Superior*. Dextra.

Hamodi, C., Moreno, J., & Barba-Martín, R. A. (2018). Medios de evaluación y desarrollo de competencias en educación superior en estudiantes de educación física. *Estudios Pedagógicos (Valdivia)*, *44*(2), 241–257. doi:10.4067/S0718-07052018000200241

Heras Bernardino, C., & Herrán Álvarez, I. (2019). La evaluación formativa y compartida desde un enfoque competencial. Aplicación práctica en tareas de aula en Primaria y Secundaria. *Revista Infancia, Educación Y Aprendizaje, 5*(2), 568–575. doi:10.22370/ieya.2019.5.2.1777

Hernando-Garijo, A., Hortigüela-Alcalá, D., & Pérez-Pueyo, Á. (2017). El proceso de evaluación formativa en la realización de un vídeo tutorial de estiramientos en inglés en un centro bilingüe. In V. M. López-Pastor & Á. Pérez-Pueyo (Eds.), *Evaluación Formativa y Compartida en Educación: experiencias de éxito en todas las etapas educativas* (pp. 260–270). Universidad de León.

Hortigüela-Alcalá, D., Abella García, V., & Pérez-Pueyo, Á. (2015). ¿ De qué manera se implica el alumnado en el aprendizaje? Análisis de su percepción en procesos de evaluación formativa. *Revista de Investigación Educacional*, *13*(1), 88–104.

Hortigüela-Alcalá, D., González-Víllora, S., & Hernando-Garijo, A. (2021). Do we really assess learning in Physical Education? Teacher's perceptions at different educational stages. *Retos*, *42*, 643–654. doi:10.47197/retos.v42i0.88686

Hortigüela-Alcalá, D., Palacios, A., & López-Pastor, V. (2019a). The impact of formative and shared or coassessment on the acquisition of transversal competences in higher education. *Assessment & Evaluation in Higher Education*, *44*(6), 933–945. doi:10.1080/02602938.2018.1530341

Hortigüela-Alcalá, D., Pérez-Pueyo, A., & González-Calvo, G. (2019b). Pero… ¿A qué nos referimos realmente con la Evaluación Formativa y Compartida?: Confusiones Habituales y Reflexiones Prácticas. *Revista Iberoamericana de Evaluación Educativa*, *12*(1), 13–27. doi:10.15366/riee2019.12.1.001

Lafuente, J. C., & Hortigüela-Alcalá, D. (2021). La percepción de los futuros maestros respecto a la implantación de contenidos de expresión corporal. *Movimento (Porto Alegre)*, *27*, 1–15. doi:10.22456/1982-8918.111735

Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students' motivation. *Assessment & Evaluation in Higher Education*, *46*(2), 236–255. doi:10.1080/02602938.2020.1765228

López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: A review of international literature. *Sport Education and Society*, *18*(1), 57–76. doi:10.1080/13573322.2012.713860

López-Pastor, V. M., & Pérez-Pueyo, Á. (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León.

MacLean, J. (2018). Teachers as Agents of Change in Curricular Reform: The Position of Dance Revisited. *Sport Education and Society*, *23*(6), 563–577. doi:10.1080/13573322.2016.1249464

Mattsson, T., & Lundvall, S. (2015). The Position of Dance in Physical Education. *Sport Education and Society*, *20*(7), 855–871. doi:10.1080/13573322.2013.837044

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

Ozan, C., & Kincal, R. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences: Theory & Practice*, 18.

Panadero, E., Fraile, J., Fernández, J., Castilla-Estévez, D., & Ruíz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, *44*(3), 379–397. doi:10.1080/02602938.2018.1512553

Pegalajar, M. del C. (2021). La Rúbrica como Instrumento para la Evaluación de Trabajos Fin de Grado. REICE. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, *19*(3), 67–81. doi:10.15366/reice2021.19.3.005

Pérez-Pueyo, A., Gutiérrez-García, C., Hortigüela-Alcalá, D., & Hernando-Garijo, A. (2017). Video-diario de evidencias de aprendizaje. Infancia. *Educación y Aprendizaje*, *3*(2), 127–132. doi:10.22370/ieya.2017.3.2.711

Pérez-Pueyo, A., Hortigüela-Alcalá, D., & Gutiérrez-García, C. (2019). La exposición oral con pecha-kucha desde la evaluación formativa. En A. Ramírez y M. P. Gutiérrez (Coords.), La evaluación educativa, entre la emoción y la razón (pp. 104-123). Universidad de Córdoba.

Romero, M., Castejón, F., López, V., & Fraile, A. (2017). Formative assessment, communicative competencies and ICT in teachers training. *Comunicar*, *52*, 1–12.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing Their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147–164. doi:10.1080/02602930301671

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*(5), 535–550. doi:10.1080/02602930903541015

Santos Guerra, M. A. (2003). Dime cómo evalúas y te diré qué tipo de profesional y de persona eres. *Revista Enfoques Educacionales*, *5*(1), 1–15. doi:10.5354/0717-3229.2003.47513

Santos Guerra, M. A. (2014). *La evaluación como aprendizaje. Cuando la flecha impacta en la diana.* Narcea.

Topping, K. J. (2009). Peer assessment. *Theory into Practice*, *48*(1), 20–27. doi:10.1080/00405840802577569

Tortorella, G., Viana, S., & Fettermann, D. (2015). Learning cycles and focus groups: A complementary approach to the A3 thinking methodology. *The Learning Organization*, *22*(4), 229–240. doi:10.1108/TLO-02-2015-0008

Weldmeskel, F. M., & Michael, D. J. (2016). The impact of formative assessment on self-regulating learning in university classrooms. *Tuning Journal for Higher Education*, *4*(1), 99–118. doi:10.18543/tjhe-4(1)-2016pp99-118

KEY TERMS AND DEFINITIONS

Assessment Criteria: They determine the degree of acquisition of key competencies and the achievement of the objectives of each teaching and educational stage.

Assessment Instrument: Documents that are linked to an assessment activity, to establish in them the aspects to be assessed, indicating the levels of achievement of each one.

Assessment Scale: Instrument in which different aspects to be assessed are shown and, in addition, each aspect is subdivided into as many aspects to be assessed, each aspect is subdivided into as many decreasing levels of achievement as it is possible to identify in the production to be assessed.

Formative Assessment: Any assessment process whose main purpose is to improve the teaching-learning processes that take place.

Grading: The process in which a numerical grade or grade is assigned to a student's learning or final production.

Shared Assessment: Assessment process that truly and continuously includes the students, understanding them as the active protagonists of their learning.

Triadic Assessment: It is the joint process of self- assessment, co- assessment and hetero-assessment with the purpose of contrasting perceptions and giving more information in the assessment processes.