

Traditional assessment and its consequences. A case in the hegemonic Physical Education

La evaluación tradicional y sus consecuencias. Un caso en la Educación Física hegemónica

Marta Oliver-Álvarez¹ 

Daniel Martos-García¹ 

¹ Facultad de Magisterio, Universidad de Valencia, Spain

Correspondence:

Marta Oliver-Álvarez
marta.oliver@uv.es

Short title:

Traditional assessment and its consequences

How to cite this article:

Oliver-Álvarez, M., & Martos-García, D. (2023). Traditional assessment and its consequences. A case in the hegemonic Physical Education. *Cultura, Ciencia y Deporte*, 18(58), 27-49. <https://doi.org/10.12800/ccd.v18i58.1976>

Received: 5 October 2022 / Accepted: 10 July 2023

Abstract

The discourse of performance in Physical Education continues to perpetuate a multitude of techniques and decisions for student assessment and grading. Physical fitness tests are commonly employed in the school context, often without due consideration of the resulting consequences. This current case study, using a qualitative methodology, aims to illustrate the experiences and interpret the consequences of using physical tests among 1st year Baccalaureate students and the respective teacher. Data collection involved a total of 14 semi-structured interviews and two-month participant observation recorded in a field diary. As the findings illustrate, the expressed emotions reveal a dichotomy between students who feel motivated by assessment tests and those who experience frustration after completing them, largely influenced by their skill level. Conversely, the teacher exhibits an urgent need to standardize the tests and shelter behind its presumed neutrality. The conclusions highlight the necessity of reflecting on the consequences of these persistently replicated practices, in order to progress towards more formative and participatory assessment processes.

Key words: Physical fitness tests, students, performance ideology, qualitative research, emotions.

Resumen

El discurso de rendimiento en Educación Física sigue reproduciendo en la evaluación un cúmulo de técnicas y decisiones para la medición y calificación del alumnado. Los test físico-deportivos se utilizan en el contexto escolar sin considerar en muchas ocasiones las consecuencias que provocan. El presente estudio de caso, a través de una metodología cualitativa, pretende mostrar las vivencias e interpretar las consecuencias que tiene la utilización de pruebas físicas para el alumnado de 1^o de Bachillerato, y su respectivo profesor. Para la recogida de datos se realizaron un total de 14 entrevistas semiestructuradas, más una observación participante de dos meses registrada en un diario de campo. Como ilustran los resultados, los sentimientos expuestos presentan una dicotomía entre el alumnado que se siente motivado hacia los test de evaluación, y el que siente frustración tras llevarlos a cabo, determinado en gran medida por su nivel de habilidad. El profesor, por su parte, muestra una necesidad imperiosa por igualar las pruebas y resguardarse en su presunta neutralidad. Las conclusiones ponen de manifiesto la necesidad de reflexionar sobre las consecuencias de las prácticas que todavía se siguen reproduciendo, para poder avanzar hacia procesos evaluativos más formativos y participativos.

Palabras clave: Test físico-deportivos, alumnado, ideología de rendimiento, investigación cualitativa, sentimientos.



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Introduction

The reconceptualization of Physical Education (PE, from now on) has been in the spotlight in recent decades. Thus, in the past, many authors have supported the idea of involving students in participatory processes and adopting a more holistic and critical approach to teaching this subject. (Fernández-Balboa, 2001; Siedentop, 1994). Decades later, contemporary authors continue to insist on this need (Crum, 2017; Kirk, 2010).

In this sense, it is fair to recognize that some modifications in terms of methodological approaches and teaching strategies have been gradually gaining ground (López-Pastor & Pérez-Pueyo, 2017). However, the changes have not reached all spheres. For instance, assessment processes are still resistant to change, even though it's true that the reforms in Spain from 1970 until now have implemented more formative and participatory assessment methods. However, the assessment system still remains, in essence, largely inflexible and slow to change (López-Pastor et al., 2013). In this regard, the most commonly used assessment in PE still relies on a combination of 'objective' and physical tests to evaluate students' physical condition. The evaluation aims to measure students' physical skills and align them with the prescribed technical model (Atienza et al., 2018; López-Pastor, 2006).

This method of evaluation responds, to a certain extent, to the needs of a subject that seeks legitimacy, and is therefore dominated by a technical discourse that has persisted for decades (Kirk, 2010; Tinning, 2017). To go into more detail, we must understand that traditional assessment methods perpetuate certain discriminatory practices, in which students are classified and differentiated between the fastest, most skilled, better ones and, on the contrary, those who are not or have inferior skills (Sánchez-Hernández et al., 2022). This approach allows teachers to clearly exercise their power, using tests to qualify and identify the most motor-successful students (Atienza et al., 2018). The main objective of the assessment is based on the students' motor performance, and so their performance and physical skills are constantly evaluated by teachers, and judged by peers (Beltrán-Carrillo & Devís, 2019). Furthermore, approaching corporeality from a mechanistic perspective overlooks the emotional and experiential aspects of movement. It fails to consider how individuals experience the proposed tasks and the meanings they create within the environment (Barbero, 2007).

Alternative positions, such as participation discourses or the critical paradigm, offer ideas and innovations that have been evolving to provide alternative methods of assessing students at different educational levels (López-Pastor & Pérez-Pueyo, 2017; McArthur, 2019). Assessment, as López-Pastor (2004) explains, is a reflection of the different ways of understanding the subject of PE, its conceptual framework, and its role in both the educational system and society. It seems necessary, therefore, to pay close attention to this subject.

Consequences of traditional and technical assessment

Despite hegemonic positions or the prevailing technical paradigm in education, which call for a more neutral assessment system, it is evident that achieving complete neutrality is simply not possible (Lorente-Catalán & Martos-García, 2018). This alludes to the fact that every decision made by teachers has its consequences, and following a traditional model is no different. In the subject of PE, for example, these decisions might harm, among other things, certain students' adherence to active lifestyles (Griggs & Fleet, 2021). In the field of evaluation, performance-based systems seek and achieve measurement, comparison, hierarchization, or control, among other things. Therefore, the effects that an objective assessment, based on achievement tests, has on students go beyond the idea of neutrality. These effects need to be explicitly addressed in order to critique them, especially if they result in discrimination.

The use, and abuse, of this type of assessment, and more specifically of physical tests, can cause negative effects on students and their subsequent relationship with physical activity (Camacho-Miñano & Prat, 2018; Danthony et al., 2020; Silverman et al., 2008). From a philosophical perspective, the desire to monitor the "risk factors" that are present in today's Western democracies has resulted in the acceptance of self-surveillance and the transformation of individuals in order to recreate the sense of security that existed in the past (Rose, 1999). Therefore, individuals should take responsibility for understanding the characteristics of these factors and strive to address them as the sole agents accountable for their decisions. One tangible risk factor in our society is the crisis of inactivity and the rise in overweight. This has led to the body being seen as an object of scrutiny, with unhealthy bodies being viewed as problematic and schools being seen as a place to enforce discipline in regard to this issue (Garrett & Wrench, 2012; Gwanas, 2002).

Given this situation, the subject of PE has been established as a tool to meet the demands of necessary physical exercise and promote healthy lifestyle habits (McKenzie, 2001; Parkinson & Burrows, 2019). Thus, the use of tests is believed to increase students' motivation to improve their physical fitness, set goals, and maintain a positive attitude to achieve those goals (Dennison et al., 1988; Welk, 2008). However, the debate around concepts such as the reliability of these tests or their ethics questions their use (Cale & Harris, 2009), given that they do not take into account the motivational, socioeconomic, or cultural factors that individuals experience (Kirk, 2006). Tests can, therefore, become a double-edged sword. While some students may find them motivating, others may have negative and humiliating experiences (Corbin, 2010; Danthony et al., 2020) that lead to disinterest in the subject, to a decrease in self-esteem, and an overall negative impact on students' self-concept (Rice, 2007).

Generating negative emotions seems clearly contrary to the supposed educational purposes of PE. Despite this, literature in our latitudes continues to uncover cases where assessments become an ordeal, as highlighted by Beltrán-Carrillo and Devís (2019), who examined the perspective of disengaged students. In this context, some publications examine the adverse effects of this type of assessment in a direct manner. They do so through quantitative studies (Danthony et al., 2020) or by analyzing the past experiences of physical education teachers in initial training (Calatayud, 2021; Valencia-Peris & Lizandra, 2018). For their part, there are narrative studies that bring to light the situations derived from practical exams as a source of fear and dread towards the PE subject (Monforte & Pérez-Samaniego, 2017).

However, it is clear that there is a lack of qualitative research on PE assessments (Atienza et al., 2018), particularly in studies that aim to understand the impact of these evaluations when they are conducted. Thus, the objective of this article is to demonstrate the experiences and analyze the outcomes of using physical tests for evaluating physical education performance in 1st year Baccalaureate students. Furthermore, the intention is to complete this interpretation with the teacher's discourse. Given this objective, we can consider the following research questions: What is the student's experience regarding their assessment? How does the teacher interpret the use of an evaluation method that relies on tests of technical execution? What are the consequences of this type of procedure?

Method

The research is situated within the interpretive paradigm and employs qualitative methodology to gain insights into the dynamics that take place within the classroom. Due to the unique characteristics of this context, this case study can serve as an important tool for generalizing the learnings (Stake, 1995) and gaining a better understanding of similar scenarios (Hodge & Sharp, 2016).

Context and participants

The study took place in a public center in the city of Valencia, specifically in a densely populated neighborhood with around 1,200 students. The socioeconomic profile of the families stated in the Center's Educational Project (PEC) corresponds to 83% middle class and 17% lower class.

Students from three different PE groups in the 1st year of Baccalaureate participated in the study. These groups were selected based on accessibility criteria, as they were the groups that the first author of this article worked with during her teaching practices, specifically during the completion of the Volleyball Didactic Unit. In total, 35 students participated, 25 girls and 10 boys, aged between 16 and 17 years. The teacher involved in the project was responsible for all three groups. He is a 56-year-old man with over 35 years of professional experience. The

evaluation included four practical exams that progressively increased in difficulty. Each exam had specific technical requirements and was assigned a certain number of points. The first exam focused on baseline serves, the second exam involved basic play including receiving, placing, and finger touch. The third exam included 2x2 play, and the fourth exam involved 4x4 play. These exams accounted for the total grade in the subject.

Procedure, material, and instruments

To collect the information, qualitative research techniques were used, specifically participant observation, document analysis and semi-structured interviews, following the guidelines established by Álvarez (2008).

The technique of participant observation was effectively balanced between observing and actively participating, based on the requirements of each situation. The first author actively engaged with the students during the execution of the motor tasks suggested by the teacher. This interaction provided her with an opportunity to observe the emotional reactions of the group members (Guba & Lincoln, 1981). However, while still keeping her research objective in mind, she incorporated moments of observation that would provide her with a different perspective on what occurred during those sessions.

The fieldwork was conducted over the two-month period when the researcher stayed at the high school, observing volleyball sessions in the sport's block. In total, the sessions consisted of approximately 54 hours of coexistence, with each group having 18 PE sessions specifically dedicated to volleyball practice. The sessions were located both in the gym and in the high school playground. Data collection was conducted using a field diary. This diary included observations on the subject's performance, informal conversations about the teacher's beliefs, and descriptions of situations during the physical tests conducted by the students.

Furthermore, there were a total of fourteen semi-structured interviews, following the principle of saturation, thirteen with the students and one with the teacher. The script was developed using existing literature on assessment and to address the intended research questions. The students who were interviewed were chosen after the first exam to include both male and female students with varying grades. This included students with very high, average, and low grades. The conversations with the students were recorded with a mobile device and transcribed verbatim. Their duration ranged from 18 to 34 minutes, and they were carried out in pairs as well as in groups of three or four people. Three of the groups consisted of four people, three consisted of three people, and seven were pairs, depending on the observed affinity between the students in the sessions. The sessions were conducted after the completion of the physical evaluation tests, before the end of the term, either in the gym or patio

of the center, depending on availability. The aim was to find a quiet place away from the rest of the class. Regarding the questions, these dealt with four fundamental topics: a) their opinions about PE (e.g., “What does the PE subject mean for you?”), b) their experiences during PE tests and exams (e.g., “How do you feel when taking the physical assessment tests?”), c) their memories of the subject (e.g., “What is your best and worst experiences related to the subject?”), and d) related questions about the fairness and unfairness of the assessment method (e.g., “What is your opinion on using physical tests to evaluate the subject?”).

For the interview with the teacher, it took place after the evaluation and after all the student interviews had been conducted. It took place in the department’s office and lasted 51 minutes. The teacher was asked about his perceptions and experiences regarding the evaluation and how it had influenced his perspective on the subject (e.g., “How would you define your way of evaluating and why do you use that method?” “To carry out the assessment, have you learned from external sources, have you read or researched about it?” “Have you had any complaints, or do you currently have any complaints regarding the type of assessment?”)

Finally, documents specific to the context were analyzed, namely the PEC and the teacher’s Classroom Program.

Data analysis and ethical and credibility criteria

Once all the interviews had been transcribed verbatim, we started delving into the data to comprehend and interpret the collected information (Levitt et al., 2018). A thorough review of the sources was conducted, and the data was then identified using codes in a Word document. This document laid the groundwork for the discussion between the two authors regarding the significance of organizing the results, clarifying how they should be arranged and grouped.

The data analysis was conducted by first reducing the amount of data and organizing it. After this, conclusions were drawn and verified (Flores, 1994). Based on the thematic analysis strategy (Braun & Clarke, 2006), we identified themes that appeared repeatedly and were of special interest for the study. This allows offering a coherent interpretation of the reality under study. Themes were subsequently grouped and synthesized into categories inductively, giving rise to four general thematic lines (metacategories) with their respective categories (Table 1).

Table 1. Thematic Analysis Lines of the Interviews

Metacategories	Category	Description
Smiles and tears, many tears	Positivity	Positive feelings towards assessment
	Satisfaction	Satisfaction after completing the tests or achieving a good grade
	Negativity	Negative feelings towards assessment
	Discomfort	Discomfort generated by the gaze of others during the tests
	Anxiety	Describing exams as strenuous test with negative health implications
	Crying	Situations of crying as a result of taking the exams
	Fear	Indicating fear as a feeling related to tests
Effort above all	Effort-result	Advocating for evaluating effort rather than just results
	Diverse abilities	Reference to the desire to set individual goals
	Injustice	Reference to assessment as an unjust practice
Tall, strong, and good	Benefited individuals	Identification of people benefitting from the assessment
	Gender differences	Identified gender differences in tests
The assessment... to learn?	Assessment concept	Perception of assessment and its utility
	Alternative assessments	Perception of other forms of assessment
	Elimination of tests	Opinion regarding the elimination of physical tests

Following Lucas and Delgado-Algarra (2020), the analysis was conducted using a multimethodical approach that allowed controlling intersubjectivity. This approach makes it possible to contrast the interpretations of the participating students with the observation data. Furthermore, the context and the participants were described following the arguments of Tracy (2010).

Regarding the ethical criteria, the students, their families, and the teacher were informed of the research aim and the three parties were given an informed consent that they signed prior to data collection. Likewise, all the names that appear are pseudonyms to avoid identification of both the students and the teacher.

Results and discussion

The results are organized into the four general thematic lines (Table 1). Their corresponding categories are outlined in each of the following sections.

Smiles and tears, many tears

As reflected in the Classroom Program, the assessment of the students consisted of four physical tests to be completed throughout the school term, ordered by increasing difficulty. These tests evoked mixed feelings among the students, ranging from pleasure to deep frustration, as previous research has shown (Hopple & Graham, 1995; Wrench, & Garrett, 2008). Students with more fit physical qualities and better results felt motivated towards this type of activities, while the rest of the students tried to avoid them (Harris & Cale, 2007). Students with successful experiences alluded to the following:

G: I think that we, being half-athletes so to speak, are not afraid of if it will go well or not, but because we have at least practiced a minimum. I at least don't think that it can go wrong as other people may think.

E: When we finished the exam, it was like, oh, that was hard, but to finish and to say, I did it! I did 30 minutes and I passed; you know?

Despite the happiness reported in the previous passages, it is still a controversial satisfaction because these tests represent examples of the "authoritarian consciousness" highlighted by Fernández-Balboa (2005, p. 130), which refers to those educationally questionable practices that the teacher, converted into an authority that distributes rewards and punishments, uncritically uses based solely on tradition. Through this repetitive mechanism, students end up accepting what happens in the classroom as something normal and even legitimate (Sánchez-Hernández et al., 2022), even though this makes them feel afraid and they even blame themselves for their actions.

This argument makes complete sense once the negative emotions and feelings that were collected during the observation are taken into account. Thus, the participating students recognized emotions such as anguish, insecurity, or anger, cataloging the practical exam as the worst moment of the subject, as could be perceived in one of the sessions: "A girl starts crying, the ball did not go over the net and she breaks down. She goes to the bathroom, and when she comes back, she says she is very distressed. I see her trembling and trying to hide".

Everything together ends up causing a feeling of incompetence among the students, a questioning of their self-concept and a clear decrease in motivation towards physical-sports activities:

C: I am frustrated by the fact that many times I feel useless for not achieving what is asked for physically, I get very frustrated because I can't do it and I have nothing that can help me solve it.

A: In the second practical exam, I ended up with an incredible anxiety attack. I couldn't finish it... I was drowning and I couldn't take it anymore.

As can be seen from this last quote, notable levels of anxiety were identified, before and during the tests. Part of the students finished the tests exhausted: "M: ... there are people who started very quickly and ended up vomiting from pushing themselves so hard".

This comment shows the level of physical demand that the tests required and, as in previous research (Huhtiniemi et al., 2021), the levels of somatic anxiety (dizziness, shortness of breath, chest tightness) that have been identified due to working close to the maximum aerobic threshold, which are similar to those presented here. Paradoxically, the purpose of these tests to promote healthy sporting practices is clearly contradicted, given how they transgress healthy limits (Cale & Harris, 2009).

Consequently, and as Rice (2007) already explained, there were negative feelings towards the subject in general. Thus, this type of experiences can lead students to avoid physical activities, something which has already been reported (Beltrán-Carrillo et al., 2012). More specifically, the assessment of PE is shown to be a source of dissatisfaction and avoidance of the subject itself. One student explained it this way:

P: It has more to do with the experience, for me it has been negative, so it clouds the fact that at the end of the day I actually had fun, but because of the bad experience, you'd rather pass on it. You look and see that it's time for PE and you go, ugh... I don't feel like it.

In this situation, the watching looks from peers and, most importantly, the teacher, become significant factors. Many times, these looks increase the discomfort of less skilled students, who perceive themselves to be the center of attention in an activity in which they do not feel competent. According to the study by Silverman et al. (2008), the affected students felt embarrassed by the way in which the tests are administered. Thus, the class could be divided, as Evans (2006) points out, between those who "watch" and those who "are watched."

Everything that happened was not hidden from the participating teacher, who chose to maintain a posture of detachment from the emotions generated by his decisions. In this way, he naturalizes the existence of students harmed by an evaluation of performance discourse:

P: Assessment, in terms of taking tests, obviously involves a series of factors that I cannot control and that are common: fear of failure, fear of being watched, fear that there is a teacher who is watching... and that causes some concern or fear in the students [...] What I try to be is as objective as possible.

His attitude demonstrated a lack of interest in self-criticism, choosing instead to rely on the objectivity of the

evidence. This raises questions about its neutrality and reopens the debate. However, recent research on fear suggests that it is usually the teachers who generate this type of emotions towards the subject, and it is in their power to mitigate this fact (Canales-Lacruz & Pina-Blanco, 2014; Monforte & Pérez-Samaniego, 2017).

Effort above all

The assessment of effort was a recurring theme in all the interviews. When faced with tests that measure their objective performance the students felt a clear desire for their effort to be valued more, rather than just focusing on the results: "P: To me, having to run for half an hour seems like something that, I don't know... that you are not being evaluated by your effort, but rather by some markers ... I don't like it."

The recurring proposal that effort be the key assessment element leads us to wonder if this idea of progress is the fairest. In this situation, it may happen that establishing effort as the only goal overlooks other learning that can and should occur in the subject (López-Pastor et al., 2016). Furthermore, despite the clear intention of the students to have a less results-oriented evaluation, the assessment of effort still remains tied to a performance ideology. This means that individuals with more appropriate physical skills and characteristics will find it easier to reach the set objectives (Valencia-Peris et al., 2020).

Thus, for example, students who do not attend extracurricular sports activities and who could benefit from positive experiences in the subject, describe their experiences in a less gratifying way than those who do attend (Säfvenbom et al., 2015):

S: Imagine that you go to basketball practice after school, then you come here and have a basketball exam, it's normal for you to get a good grade, but what I'm referring to is that a person who makes an effort and who doesn't have afterschool basketball practice also deserves that grade.

This perspective places PE in line with the neoliberal discourses that govern today's society, where the search for efficiency prevails and in which students with a greater 'innate' capacity surpass their peers and enjoy their achievements. In this sense, establishing effort as a core element of assessment does not take into account the motivational and resource differences that exist between students and that can condition their situation (Evans, 2014).

In a certain way, a PE subject develops in which there is a clear group of students harmed by the technical orientation of the assessment. This, among other things, can generate certain conflicts, for example, when disputing their grades. Faced with this possibility, the participating teacher hid behind the objectivity of his method:

P: Experience has told me that everything related to subjectivity causes problems, in the sense that there will always be someone who feels unfairly evaluated. When

confronted with an objective test, transparency means that possible repercussions or complaints are minimized in the face of something that is neutral, and that can be the same for all.

And in relation to what is fair and what is unfair, he argued: "P: I think it is unfair. But it is what gives me greater peace of mind when evaluating."

Without a doubt, the comfort and peace of mind that this type of assessment provides to teachers continues to be seen as one of its greatest advantages. Sometimes, exploring the unknown can be compromised by the lack of training and resources, which prevents teachers from a more formative assessment method (López-Pastor et al., 2016). In this specific case, the weight of tradition and the previous background that the teacher possessed deeply marked his vision of assessment, limiting the possibilities towards alternative methods.

Tall, strong, and good

Under this paradigm, the differences in the individual skills of students are perceived as the traits that determine which people benefit and which do not. This is ableism (Wolbring, 2008) in all its expression. A student referred to this:

L: I know how to do the technique like this, more or less, but if I don't have the strength, or I have never played this sport, or I don't have the skill, then... compared to other people, I am inferior.

Thus, in the developed exams and tests, physical qualities according to performance standards allow arranging the students hierarchically, as described by Fernández-Balboa and Muros (2006). It is not surprising then that, similar to the study by Beltrán-Carrillo et al. (2012), students perceived that PE classes were organized around activities that only the most skilled could achieve, using the latter as a measure to judge the rest of the class:

L: And I don't like it when the good ones are on one team and the bad ones on the other. He always says, you make the teams, I don't make them... so then he is allowing the good ones to get good grades and the bad ones to just get frustrated among ourselves.

As confirmed by Aggerholm et al. (2018), it is common for a hierarchy to form during volleyball tests in the class, dividing the skilled and successful students from those who were not as proficient or successful: "The students are in the gym practicing on their own for the exam. On one side are the most skilled kids performing the exercises without difficulty, while on the other side there are mixed groups with more motor problems." This way of separating the students, based on a sporting ideal, tends to marginalize, and differentiate students based on their psychomotor level, creating what Barbero (1996, p. 28) calls "pockets of clumsiness". This supports the criticism that traditional evaluation receives in our field due to its competitive format, which promotes constant comparison among students.

In this sense, the confirmation of a body ideal, which takes as a reference the 'tall, strong and good', is conflictive, especially among the students: "The teacher first asks if anyone wants to start with 2x2 competing. Three boys raise their hands. The girls whisper among themselves: they know how to hit the ball, position themselves, play, we don't".

As the literature shows, low perceived competence is presented as an element that conditions enjoyment in PE classes (Evans, 2006). And given that the curriculum is usually overrepresented by sports activities in which skills with a masculine cultural background predominate (Monforte & Úbeda-Colomer, 2019), the standards of success are then also based on male students. Consequently, the female students feel less capable and competent, which conditions their motivation towards the subject in general and towards the assessment tests in particular (Garrett, 2004; Simonton et al., 2019).

Female students' participation problems in PE are often attributed to a lack of aptitude that causes an absence of attitude (Martos-García et al., 2020), in other words, girls are seen 'as a problem'. The participating teacher is no exception and, as he argued in an informal conversation, "he attributed the differences in participation to a previous sports culture among male students, which female students lacked, and he convinced himself that he was incapable of solving this problem, blaming it on sociocultural factors that he was incapable of changing". Some comments from female students reinforced this fact, assuming a position of inferiority with regard to boys: "C: Only because of how good they all are, the boys, taller, more everything... they stand out a lot because they play very well, and well, you feel like, I am so bad, so clumsy."

All in all, given a results-oriented vision of the subject that seeks efficient bodies, it is not surprising that hegemonic masculine values such as roughness, strength, or vigor are frequently exalted (Stride & Flintoff, 2018). This exclusion, with motor skills as a justification, not only affects female students but also affects those students who do not meet the male ideal:

E: In the end I also think that the boys feel a lot more pressure, because their goals to achieve are higher than ours... For example, Pablo or Dani, do not have the same physical condition as the good ones. So, having the same goals, all boys are not equal, and demanding the same from all of them is just not fair.

As we see, gender alone does not always explain these facts. According to Nielsen and Pfister (2011), there is a percentage of boys who don't conform to the socially established male standards and don't adhere to the dominant hegemonic idea of masculinity that is highlighted in the context of physical education. Therefore, when the subject is permeated with a competitive atmosphere that rewards certain physical conditions, it becomes a place of privilege and power for one sector of the student body,

and of rejection and avoidance for another. A space where individuals should be able to enjoy a wide range of fulfilling experiences is marred by instances of discrimination.

The assessment... To learn?

The data show a traditional assessment system based on practical tests from which the grade is directly derived. The teacher undoubtedly contributed to feeding the students' obsession with grades, which depended to a greater extent on motor skills:

S: I relate it a little to your physical level, right? I don't know... it's just that they evaluate you according to your skills.

F: At the end of the day, it is an exam, you do the best you can, and you have the grade for what you do and what you are.

As mentioned in previous comments, the subject is associated with a training perspective that aligns with technical rationality and hegemonic discourse. This perspective prioritizes the student's performance as the ultimate objective, focusing on achieving good grades, "given how it measures what you are, instead of the actual learning of the content taught" (López-Pastor et al., 2006, p. 34).

In this sense, adopting alternative assessment systems, in which the processes focus on learning and not so much on the result, is a possibility. However, introducing alternative forms of evaluation does not simply lie in exchanging evaluation instruments, but in reestablishing power relationships in the classroom and "redefining what it means to be the 'authority' or 'expert' in a given field" (Lorente - Catalán & Kirk, 2013, p. 78). It does not seem, therefore, that the teacher's conception of PE and assessment is in line with these postulates, and this reinforces the fact that teachers become one of the main obstacles to change in terms of assessment (Hamodí et al., 2017; Martínez-Rizo, 2013).

This inclination exposes the prevalence of exam culture in the educational system and how it leads to the emergence of an authoritarian mindset, as discussed earlier. It compels students to accept and consider these tests as normal, despite their discriminatory nature (Fernández-Balboa, 2005).

Conclusions

The traditional and hegemonic conception of education has left its mark on the way of understanding assessment in the subject of PE, giving rise to inherited practices far from the formative approaches currently pursued (Blázquez, 2017).

Taking into consideration the objectives outlined in this study, we have explored the students' interpretations of an evaluation that focuses on technical aspects, delving into the emotional and social aspects that contributed to shaping their experiences. While a part of the students

could enjoy or feel motivated with this type of assessment, for the vast majority it triggered feelings that ranged from nervousness, anxiety, frustration, or fear, and even developed a concept of inability towards their own skills. In this sense, the tests reinforced the skills and supremacy of the majority of the boys given that, in general, they were the only ones who had access to the highest grades. On the other hand, for many girls and some boys they served as a source of dissatisfaction, through which they came to assume and normalize an inferior role that, among other things, led to aversion towards the subject.

With the results presented, the supposed neutrality of the technical tests is called into question. These tests were intended to promote healthy lifestyle habits, but they only seem to benefit students who already possess the necessary characteristics. In addition, these tests do not seem to take into account economic, social, or motivational factors that can influence students' well-being and their commitment to physical activity (Wrench & Garret, 2008). Furthermore, the relationship between evaluation and learning in this type of approach is almost non-existent. The students' perceptions revealed the purely qualifying usefulness of the assessment, given that they only felt that it served to measure and judge them based on their physical skills.

In the same way, the vision of the teacher has been put on the table, his comfort and professional security being among the reasons for promoting performance discourse. In his intent to equalize the tests, he promoted the differences that the students already presented, creating an even more marked hierarchy in the sessions. Despite recognizing these issues, the peace of mind that the presumed objectivity of the tests gave him made it impossible for him to explore another type of assessment that would entail a more personal and qualitative involvement.

In this sense, it seems pertinent to continue interpreting the reasons and barriers that prevent PE teachers from developing more participatory and formative assessment processes, as encouraged by MacPhail and Murphy (2017) for the Irish context. This is just one case, which represents a specific discourse and its consequences, which should continue to be explored.

Among the main limitations of the study, we highlight the shortness of the participant observation time, being only two months, which could undermine a deeper understanding of the reality in question. In addition, interviews could have been carried out with a larger number of students, even from other grades and groups, with the aim of delving into the consequences generated by this type of assessment.

Finally, considering the impact of this evaluation approach, it is important to conduct studies that apply innovative evaluation designs, aligned with the discourses of participation and criticism. These designs should offer

motor tasks and strategies that assist in addressing any differences in skills among students and help in their reconstruction. Fortunately, there are a number of alternative proposals with successful results that have been developed using a more formative and participatory assessment method in different educational stages (Asún et al., 2017; Fernández et al., 2019; López-Pastor & López-Luengo, 2005; López-Pastor & Pérez-Pueyo, 2017; Lorente-Catalán & Kirk, 2013; Molina-Soria et al. 2023). It would be advisable to extend these and further evaluate them. We propose, therefore, to move towards more democratic educational paradigms, away from the traditional and technical paradigm, reinforcing the purely educational orientation of the subject of PE.

Bibliography

- Aggerholm, K., Standal, Ø. F., & Hordvik, M. M. (2018). Competition in physical education: Avoid, ask, adapt or accept?. *Quest*, 70(3), 385-400. <https://doi.org/10.1080/00336297.2017.1415151>
- Álvarez, C. (2008). La etnografía como modelo de investigación en educación. *Gazeta de Antropología*, 24(1), 1-15. <https://doi.org/10.30827/Digibug.6998>
- Asún, S., Romero, M. R., Aparicio, J. L., & Fraile, A. (2017). Evaluación formativa en Expresión Corporal. *TANDEM. Didáctica de la EF. Monográfico invitación a la danza*, (55), 38-45. <https://dialnet.unirioja.es/servlet/articulo?codigo=5900355>
- Atienza, R., Valencia, A., & Devís, J. (2018). Experiencias de Evaluación en Educación Física. Una Aproximación desde la Formación Inicial del Profesorado. *Estudios Pedagógicos (Valdivia)*, 44(2), 127-147. <https://doi.org/10.4067/S0718-07052018000200127>
- Barbero, J. I. (1996). Cultura profesional y currículum (oculto) en educación física. Reflexiones sobre las (im)posibilidades del cambio. *Revista de Educación*, (331), 13-49. <https://dialnet.unirioja.es/servlet/articulo?codigo=19107>
- Barbero, J. I. (2007). Capital (es) corporal (es) que configuran las corrientes y/o contenidos de la educación física escolar. *Ágora para la Educación Física y el Deporte*, 4-5, 21-38. <https://dialnet.unirioja.es/servlet/articulo?codigo=2524857>
- Beltrán-Carrillo, V. J., & 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. <https://doi.org/10.5232/ricyde2019.05502>
- Beltrán-Carrillo, V. J., Devís, J., Peiró-Velert, C., & Brown, D. H. (2012). When physical activity participation promotes inactivity: Negative experiences of Spanish adolescents in physical education and

- sport. *Youth & Society*, 44(1), 3-27. <https://doi.org/10.1177/0044118X10388262>
- Blázquez, D. (2017). *Cómo evaluar bien en Educación Física. El enfoque de la evaluación formativa*. Inde.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Calatayud, M. (2021). Metáforas evaluativas del alumnado universitario de Educación Física. Necesidad de un cambio radical. *Retos*, 41, 854-865. <https://doi.org/10.47197/retos.v41i0.82706>
- Cale, L., & Harris, J. (2009). Fitness testing in physical education—a misdirected effort in promoting healthy lifestyles and physical activity? *Physical Education and Sport Pedagogy*, 14(1), 89-108. <https://doi.org/10.1080/17408980701345782>
- Camacho-Miñano, M. J., & Prat, M. (2018). Violencia simbólica en la educación física escolar: un análisis crítico de las experiencias negativas del futuro profesorado de educación primaria. *Movimiento*, 24(3), 815-826. <https://doi.org/10.22456/1982-8918.79171>
- Canales-Lacruz, I., & Pina-Blanco, I. (2014). El miedo al contacto en el voleibol. Percepciones del alumnado de Educación Física. *Ágora para la Educación Física y el Deporte*, 16(2), 122-136. <https://uvadoc.uva.es/bitstream/handle/10324/23774/AEFD-2014-2-miedo-contacto-voleibol.pdf?sequence=1&isAllowed=y>
- Corbin, C. (2010). Texas youth fitness study: a commentary. *Research quarterly for exercise and sport*, 81(sup3), S75-S78. <https://doi.org/10.1080/02701367.2010.10599696>
- Crum, B. (2017). How to win the battle for survival as a school subject? Reflections on justification, objectives, methods and organization of PE in schools of the 21st century. *RETOS-Nuevas Tendencias en Educación Física, Deporte y Recreación*, 31, 238-244. <https://doi.org/10.47197/retos.v0i31.53496>
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. Jossey-Bass.
- Danthonny, S., Mascaret, N., & Cury, F. (2020). Test anxiety in physical education: The predictive role of gender, age, and implicit theories of athletic ability. *European Physical Education Review*, 26(1), 128-143. <https://doi.org/10.1177/1356336X19839408>
- Dennison, B. A., Straus, J. H., Mellits, E. D., & Charney, E. (1988). Childhood physical fitness tests: predictor of adult physical activity levels. *Pediatrics*, 82(3), 324-330. <https://doi.org/10.1542/peds.82.3.324>
- Evans, B. (2006). 'I'd feel ashamed': Girls' bodies and sports participation. *Gender, place & culture*, 13(5), 547-561. <https://doi.org/10.1080/09663690600858952>
- Evans, J. (2014). Neoliberalism and the future for a socio-educative physical education. *Physical Education and Sport Pedagogy*, 19(5), 545-558. <https://doi.org/10.1080/17408989.2013.817010>
- Fernández-Balboa, J. M. (2001). La sociedad, la escuela y la educación física del futuro. En J. Devís (Coord.), *La Educación Física, el deporte y la salud en el siglo XXI* (pp. 25-45). Marfil.
- Fernández-Balboa, J. M. (2005). La auto-evaluación como práctica promotora de la democracia y la dignidad. En Á. S. Camacho & J. M. Fernández-Balboa (Eds.), *La otra cara de la enseñanza: La educación física desde una perspectiva crítica* (pp. 127-158). INDE Publicaciones.
- Fernández-Balboa, J. M., & Muros, B. (2006). The hegemonic triumvirate—Ideologies, discourses, and habitus in sport and physical education: Implications and suggestions. *Quest*, 58(2), 197-221. <https://doi.org/10.1080/00336297.2006.10491879>
- Fernández, C. López-Pastor, V. M., & Pascual, C. (2019). Aprendiendo a desarrollar sistemas de evaluación formativa y compartida en las prácticas como maestra de Educación Física en Primaria. *Revista de Innovación y Buenas Prácticas Docentes*, 8(1), 119-131. <https://journals.uco.es/index.php/ripadoc/article/view/12001>
- Flores, J. G. (1994). *Análisis de datos cualitativos: aplicaciones a la investigación educativa*. Promociones y Publicaciones Universitarias, PPU, S.A.
- Garrett, R. (2004). Negotiating a physical identity: Girls, bodies and physical education. *Sport, education and society*, 9(2), 223-237. <https://doi.org/10.1080/1357332042000233958>
- Garrett, R., & Wrench, A. (2012). 'Society has taught us to judge': Cultures of the body in teacher education. *Asia-Pacific Journal of Teacher Education*, 40(2), 111-126. <https://doi.org/10.1080/1359866X.2012.669826>
- Griggs, G., & Fleet, M. (2021). Most people hate physical education and most drop out of physical activity: In search of credible curriculum alternatives. *Education Sciences*, 11(11), 701. <https://doi.org/10.3390/educsci11110701>
- Gwanas, B. (2002). *Constructing body space: gender, sport and body image in adolescence*. [Doctoral dissertation, University of Liverpool]. ETHOS British Library. <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.288204>
- Hamodi, C., López-Pastor, V. M., & López-Pastor, A. T. (2017). If I experience formative assessment whilst studying at university, will I put it into practice later as a teacher? Formative and shared assessment in Initial Teacher Education (ITE). *European Journal of Teacher Education*, 40(2), 171-190. <https://doi.org/10.1080/02619768.2017.1281909>

- Harris, J., & Cale, L. (2007). Children's fitness testing: A feasibility study. *Health Education Journal*, 66(2), 153-172. <https://doi.org/10.1177/0017896907076754>
- Hodge, K., & Sharp, L. A. (2016). Case studies. En B. Smith & A. C. Sparkes (Eds.), *Handbook of Qualitative Research in Sport and Exercise* (pp. 62-74). Routledge
- Hopple, C., & Graham, G. (1995). What children think, feel, and know about physical fitness testing. *Journal of Teaching in Physical Education*, 14(4), 408-417. <https://doi.org/10.1123/jtpe.14.4.408>
- Huhtiniemi, M., Salin, K., Lahti, J., Saakslahki, A., Tolvanen, A., Watt, A., & Jaakkola, T. (2021). Finnish student's enjoyment and anxiety levels during fitness testing classes. *Physical Education and Sport Pedagogy*, 26(1), 1-15. <https://doi.org/10.1080/17408989.2020.1793926>
- Kirk, D. (2006) 'The "Obesity Crisis" and School Physical Education'. *Sport Education and Society*, 11, 121-33. <https://doi.org/10.1080/13573320600640660>
- Kirk, D. (2010). *Physical education futures*. Routledge. <https://doi.org/10.4324/9780203874622>
- Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suarez-Orozco, C. (2018). Journal article reporting standards for qualitative research in psychology: The APA publications and communications board task force report. *American Psychologist*, 73(1), 26-46. <https://doi.org/10.1037/amp0000151>
- López-Pastor, V. M. (2006). *La Evaluación en Educación Física: revisión de los modelos tradicionales y planteamiento de una alternativa: la evaluación formativa y compartida*. Miño y Dávila.
- López-Pastor, V. M., Aguado, R. M., García, J. G., Pastor, E. M. L., Pinela, J. F. M., Badiola, J. G., Barba-Martín, J. J., Aguilar-Baeza, R., González-Pascual, M., Heras-Bernardino, C., Martín, M. I., Manrique-Arribas, J. C., Subtil-Maraguán, P. & Maraguán-García, L. (2006). La evaluación en educación física. Revisión de modelos tradicionales y planteamiento de una alternativa. La evaluación formativa y compartida. *Retos. Nuevas tendencias en Educación Física, Deporte y Recreación*, 10, 31-41. <https://doi.org/10.47197/retos.v0i10.35061>
- López-Pastor, V. M., & López-Luengo, M. (21-23 de septiembre de 2005). *Evaluación, calificación e innovación en la enseñanza universitaria: la evaluación formativa y compartida: dificultades, 42 posibilidades, encuentros y desencuentros* [Comunicación]. II Jornadas Internacionales De Innovación Universitaria, Universidad Europea de Madrid.
- López-Pastor, V. M., Pérez-Brunicardi, D., Manrique, J. C., & Monjas, R. (2016). Los retos de la Educación Física en el Siglo XXI. *Retos. Nuevas tendencias en Educación Física, deporte y recreación*, 29, 182-187. <https://doi.org/10.47197/retos.v0i29.42552>
- 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.
- 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. <https://doi.org/10.1080/13573322.2012.713860>
- Lorente-Catalán, E., & Kirk, D. (2013). Alternative democratic assessment in PETE: an action research study exploring risks, challenges and solutions. *Sport, Education and Society*, 18(1), 77-96. <https://doi.org/10.1080/13573322.2012.713859>
- Lucas, L., & Delgado-Algarra, E. J. (2020). El profesor posmoderno de Ciencias Sociales: Un modelo de buenas prácticas en educación patrimonial. *REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 18(1), 27-45. <https://doi.org/10.15366/reice2020.18.1.002>
- Lucea, J. D. (2005). *La evaluación formativa como instrumento de aprendizaje en Educación Física*. Inde.
- MacPhail, A., & Murphy, F. (2017). Too much freedom and autonomy in the enactment of assessment? Assessment in physical education in Ireland. *Irish Educational Studies*, 36(2), 237-252. <https://doi.org/10.1080/03323315.2017.1327365>
- Martos-García, D., Fernández-Lasa, U., & Usabiaga, O. (2020). Coeducación y deportes colectivos. La participación de las alumnas en entredicho. *Cultura, Ciencia y Deporte*, 14(4), 411-419. <https://doi.org/10.12800/ccd.v15i45.1518>
- Martos-García, D., Lorente-Catalán, E. & Martínez-Bonafé, J. (2018). Educación Física y Pedagogía Crítica: una necesidad educativa. En E. Lorente-Catalán y D. Martos-García (Eds.), *Educación Física y pedagogía crítica. Propuestas para la transformación personal y social* (pp. 29-52). Edicions de la Universitat de Lleida/ Publicacions de la Universitat de València.
- Martínez-Rizo, F. (2013). Dificultades para implementar la evaluación formativa: Revisión de literatura. *Perfiles educativos*, 35(139), 128-150. <https://doi.org/10.22201/iisue.24486167e.2013.139.35716>
- McArthur, J. (2019). *Assessment for social justice: Perspectives and practices within higher education*. Bloomsbury Publishing.
- McKenzie, G. (2001). Physical activity and health: School interventions. *Abstracts of the 6th Annual Congress of the European College of Sports Science*, 17, 24-28.
- Molina-Soria, M., López-Pastor, V. M., Hortigüela-Alcalá, D., Pascual-Arias, C., & Fernández-Garcimartín, C. (2023). Formative and Shared Assessment and

- Feedback: an example of good practice in Physical Education in Pre-service Teacher Education. *Cultura, Ciencia y Deporte*, 18(55). <https://doi.org/10.12800/ccd.v18i55.1986>
- Monforte, J., & Pérez-Samaniego, V. (2017). El miedo en educación física: Una historia reconocible. *Movimiento (ESEFID/UFRGS)*, 23(1), 85-100. <https://doi.org/10.22456/1982-8918.71272>
- Monforte, J., & Úbeda-Colomer, J. (2019). 'Como una chica': un estudio provocativo sobre estereotipos de género en educación física. *Retos: Nuevas tendencias en Educación Física, Deportes y Recreación*, 36, 74-79. <https://doi.org/10.47197/retos.v36i36.68598>
- Parkinson, S., & Burrows, A. (2020). Physical educator and/or health promoter? Constructing 'healthiness' and embodying a 'healthy role model' in secondary school physical education. *Sport, Education and Society*, 25(4), 365-377. <https://doi.org/10.1080/13573322.2019.1613635>
- Rice, C. (2007). Becoming "the fat girl": Acquisition of an unfit identity. *Women's Studies International Forum*, 30(2), 158-174. <https://doi.org/10.1016/j.wsif.2007.01.001>
- Rose, N. (1999). *Powers of freedom: Reframing political thought*. Cambridge university press. <https://doi.org/10.1017/CBO9780511488856>
- Säfvenbom, R., Haugen, T., & Bulie, M. (2015). Attitudes toward and motivation for PE. Who collects the Benefits of the subject? *Physical education and sport pedagogy*, 20(6), 629-646. <https://doi.org/10.1080/17408989.2014.892063>
- Sánchez-Hernández, N., Soler-Prat, S. & Martos-García, D. (2022). La Educación Física desde dentro. El discurso del rendimiento, el currículum oculto y las discriminaciones de género. *Ágora para la Educación Física y el Deporte*, (24), 46-71 <https://doi.org/10.24197/aefd.24.2022.46-71>
- Siedentop, D. (1994). *Sport education: Quality PE through positive sport experiences*. Human Kinetics
- Silverman, S., Keating, X. D., & Phillips, S. R. (2008). A lasting impression: A pedagogical perspective on youth fitness testing. *Measurement in Physical Education and Exercise Science*, 12(3), 146-166. <https://doi.org/10.1080/10913670802216122>
- Simonton, K. L., Mercier, K., & Garn, A. C. (2019). Do fitness test performances predict students' attitudes and emotions toward physical education? *Physical Education and Sport Pedagogy*, 24(6), 549-564. <https://doi.org/10.1080/17408989.2019.1628932>
- Stake, R.E. (1995). *The Art of Case Study*. Sage.
- Stride, A., & Flintoff, A. (2018). Girls, physical education and feminist praxis. En L. Mansfield, J. Caudwell, B. Wheaton, & B. Watson (Eds.), *The Palgrave Handbook of Feminism and Sport, Leisure and Physical Education* (pp. 855-869). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-53318-0_54
- Tinning, R. (2017). Transformative pedagogies and physical education. En C. D. Ennis (Ed.), *Routledge Handbook of Physical Education Pedagogies* (pp. 281-294). Routledge.
- Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative inquiry*, 16(10), 837-851. <https://doi.org/10.1177/1077800410383121>
- Valencia-Peris, A., & Lizandra, J. (2018). Cambios en la representación social de la educación física en la formación inicial del profesorado. *Retos. Nuevas tendencias en Educación Física, Deportes y Recreación*, 34, 230-235. <https://doi.org/10.47197/retos.v0i34.60144>
- Valencia-Peris, A., Salinas-Camacho, J., & Martos-García, D. (2020). Currículum oculto en educación física: un estudio de caso. *Apunts. Educación física y deportes*, 3(141), 33- 40. [https://doi.org/10.5672/apunts.2014-0983.es.\(2020/3\).141.04](https://doi.org/10.5672/apunts.2014-0983.es.(2020/3).141.04)
- Welk, G.J. (2008). The Role of Physical Activity Assessments for School-Based Physical Activity Promotion. *Measurement in Physical Education and Exercise Science*, 12(3), 184-206. <https://doi.org/10.1080/10913670802216130>
- With-Nielsen, N., & Pfister, G. (2011) Gender constructions and negotiations in physical education: Case studies. *Sport, Education and Society*, 16(5), 645-664. <https://doi.org/10.1080/13573322.2011.601145>