

1 **Challenges with using Formative and Authentic**
2 **Assessment in Physical Education teaching from**
3 **experienced teachers' perspectives**
4

5 Emilio J. Barrientos Hernán

6 Facultad de Educación de Segovia, Universidad de Valladolid, Madrid, Spain

7 <https://orcid.org/0000-0002-6060-8486>
8

9 Víctor M. López-Pastor

10 Facultad de Educación de Segovia, Departamento de Didáctica de la Expresión Musical,
11 Plástica y Corporal, Universidad de Valladolid, Valladolid, Spain

12 <https://orcid.org/0000-0003-2681-9543>
13

14 Eloísa Lorente-Catalán

15 Institut Nacional d'Educació Física de Catalunya (INEFC), Universitat de Lleida, Campus de
16 Lleida, Lleida, Spain

17 <https://orcid.org/0000-0003-3244-8013>
18

19 David Kirk

20 School of Education, University of Strathclyde, Strathclyde, UK

21 <https://orcid.org/0000-0001-9884-9106>
22

1 **Abstract**

2 **Purpose:** The main purpose of this study was to identify the challenges with using FA and
3 AA in PE teaching, based on some experienced teachers' perspectives.

4 **Method:** Semi-structured interviews to four teachers who were experienced in the area of FA
5 and AA were conducted and analysed.

6 **Results:** Revealed how the teachers operationalized FA and AA to support student learning.
7 They emphasized that to obtain positive results for teachers and pupils, it is key that the
8 teaching is organised with assessment in mind, and that the assessment is structured.

9 **Conclusions:** With pedagogic knowledge and positive attitude to use FA and AA it is possible
10 to overcome the drawbacks revealed in the PE assessment literature. Nevertheless, these
11 teachers belonged to communities of practice linked to FA and AA in PE which seems to be a
12 key factor to achieve the desired "assessment literacy" needed to initiate those changes in PE.

13

14 **Keywords:** Self-Assessment, Teacher Education, Peer-Assessment, Assessment for Learning.

15

1 **Introduction**

2 Recently, the International Association for Physical Education in Higher Education advocated
3 for the importance of assessment practices as central to providing meaningful, relevant and
4 worthwhile physical education (AIESEP, 2020, p.2). Nevertheless, assessment has been one of
5 the most problematic issues to deal with in Physical Education (PE) over the last 40 years
6 (López-Pastor et al., 2013). Traditionally, the predominant type of assessment in PE at the
7 compulsory stage of school education has been characterized as unsystematic and unstructured
8 (López-Pastor et al., 2013). PE teachers carry out direct observation of pupil behaviour,
9 sometimes employing criteria but often not. They do not usually record that information using
10 concrete assessment instruments, and pupils are typically not involved in the assessment
11 process (Dinan-Thompson & Penney, 2015; Mintah, 2003). The research literature also
12 suggests that for the compulsory years of schooling PE teachers do not systematically generate
13 information about pupil learning upon which to base their teaching. Teachers often cite
14 workload as a reason for this (Dinan-Thompson & Penney, 2015; Mintah, 2003). Therefore, it
15 seems that assessment for learning (AfL) formative assessment (FA) theory is not being applied
16 to practice in the classroom (Leirhaug & MacPhail, 2015; Leirhaug & Annerstedt, 2016;
17 Borghouts et al., 2016). PE teachers know that they should employ formative assessment, but
18 they can't because the learning objects are for the most part unarticulated (Borghouts et al.,
19 2016; Nyberg & Larsson, 2014).

20 In Spain, the country in which this study is situated, the orientation given in the national
21 syllabus is that the assessment of students' learning will be continuous and will take into
22 account their progress in all areas. Teachers have freedom to design the assessment process
23 and only have the obligation to grade pupils every term of the academic year. For that reason,
24 although assessment practices differ among teachers, the same problems found in the
25 International literature arise in the Spanish context.

1 An enduring challenge, then, is for teachers to engage in quality assessment practices,
2 considering those that are not onerous to implement and that provide good quality information
3 teachers and pupils can use to progress pupil learning (Redelius & Hay, 2009). However,
4 according to AIESEP position statement on assessment, conducting good quality assessment
5 practices requires teachers to develop assessment literacy (AIESEP, 2020). This international
6 association has defined assessment literacy as the set of beliefs, knowledge and practices about
7 assessment that lead a teacher to use assessment to improve student learning and achievement
8 (AIESEP, 2020, p.5).

9 In this context, it would be interesting to know what some PE teachers experienced in the use
10 of assessment say about how they apply these ideas in the classroom and what challenges they
11 meet. Specifically, the research objectives are: a) to analyse how a purposive sample of
12 assessment literate in-service PE teachers described their challenges with using FA and
13 Authentic Assessment (AA) and b) to analyse how these teachers recorded FA and AA
14 evidence. The challenges encountered by them can be useful to other practitioners who want
15 to initiate these changes in their teaching and may help to redirect teacher training.

16 Next, we will define the concepts of FA and AA and track the evolution of the concept of FA
17 in PE and how it has evolved towards new ways of defining and understanding assessment.

18 ***Formative Assessment and Authentic Assessment in Physical Education***

19 When assessment in PE is considered and implemented, there seem to be opposing models. On
20 the one hand, we have systems that approximate the 'traditional-summative' approach, while
21 on the other hand we have the systems that represent the 'alternative forms' approach. The
22 traditional-summative assessment approach in PE usually consists of assessment systems
23 focused on how to arrive at the grade/qualification of the pupils at the end of a term or course
24 (Slingerland et al., 2017), in which physical fitness tests and/or motor skill tests predominate.
25 In these cases, the pupil's final grade, or at least a significant percentage of it, derives from the

1 results obtained in these tests (Desrosiers, Godbout, & Genet-Volet, 1997; MacPhail &
2 Murphy, 2017). On the other hand, 'alternative assessments' often involve students in actively
3 solving realistic problems through application of new information, prior knowledge and
4 relevant skills and teachers using this information to inform their teaching and support pupils'
5 learning (López-Pastor et al., 2013; Siedentop & Tannehill, 2000).

6 It is known that assessment has different functions in the educational process: diagnostic,
7 formative and summative. In this sense the AIESEP position statement differentiates between
8 the concepts of "AfL" (related to formative assessment) and "Assessment of Learning" (AoL,
9 related to summative assessment) (AIESEP, 2020). Although these concepts are
10 complimentary (Scanlon, MacPhail, Walsh, & Tannehill, 2022) a key difference between these
11 approaches is not so much what is measured and how, but how the data generated through
12 feedback and the means of generating it, are used pedagogically. In the AfL, pupils are more
13 likely to be actively involved in the assessment process as well as using information to better
14 understand and improve their own learning. In this study, we will focus on teachers who have
15 already made this conceptual shift and use alternative assessment approaches such as FA and
16 AA. According to López-Pastor et al. (2013, p.67) FA is the assessment process aimed at
17 enhancing the teaching-learning procedures taking place. Any assessment process which helps:
18 (1) students to learn more and correct their own mistakes; (2) teachers to learn to enhance their
19 teaching practice and; (3) subject or programme development to run at its best. The purpose of
20 this alternative form of assessment is not to 'sift and sort' pupils, but to obtain information that
21 helps pupils to improve and learn and to have positive impact on the pupils' and teachers'
22 perceptions about PE (MacPhail & Murphy, 2017; Redelius & Hay, 2009; Tolgfors & Öhman,
23 2015). They recognise, too, that assessment must be a dialogical, mutual and collective
24 decision-making process among and between pupils and their teacher rather than an individual
25 and imposed process. The techniques for involving pupils in assessment are usually through

1 self-assessment, peer-assessment and co-assessment (Lorente-Catalán & Kirk, 2014a).
2 According to Hay, (2006); López-Pastor et al., (2013); MacPhail & Murphy, (2017); Penney,
3 Broker, Hay & Gillespie, (2009); Redelius & Hay, (2009), FA appears to improve the quality
4 of teaching and learning processes in PE and also has a positive impact on the perception of
5 pupils and teachers about PE (López-Pastor, 2006a; Ní Chróinín & Cosgrave, 2013). In
6 addition, the fact that pupils are participants in the creation of assessment criteria for their own
7 learning helps them learn more and be more involved in the teaching and learning processes
8 (Svennberg, Meckbach & Redelius, 2014).

9 If we analyse the international literature on FA in PE, we see that the term FA has evolved and
10 other terms have emerged (López-Pastor et al., 2013; Richard & Godbout, 2000), such as
11 Assessment for Learning (AfL) (Casbon & Spackman, 2005; Chng & Lund, 2018; Hay, 2006).
12 In the last decade, the term most often linked to FA seems to be AfL (Borghouts, Slingerland,
13 & Haerens, 2016; Hay, 2006; Leirhaug, 2016; Leirhaug & MacPhail, 2015; López-Pastor et
14 al., 2013; Tolgfors & Öhman, 2015).

15 The concept of AfL can be characterized by five principles according to Leirhaug and MacPhail
16 (2015) and Tolgfors and Öhman (2015). These are: clarifying learning intentions and criteria
17 for success; engineering effective classroom discussions and other learning tasks that elicit
18 evidence of pupil understanding; providing feedback that moves learners forward; activating
19 pupils as instructional resources for one another; activating pupils as the owners of their own
20 learning.

21 Leirhaug and MacPhail (2015) carried out a study in Norway with three PE teachers to identify
22 what type of assessment they were using and if it constituted AfL. They discovered that the
23 assessment practices of these teachers fell short of meeting the five principles mentioned above
24 for them to be considered AfL. In a similar vein, Leirhaug and Annerstedt (2016) investigated
25 whether a curriculum focused on AfL was being used in PE classes in Norway. Pupils pointed

1 out in this case that PE practices did not reflect the principles of AfL and that teachers'
2 understanding of these principles may vary depending on how they interpret them. This could
3 indicate that, although educational authorities have introduced AfL within policy, it is not being
4 applied in PE classes in a satisfactory manner. On the other hand, Borghouts et al. (2016)
5 conducted a study in Holland with 206 PE teachers to identify the importance they attached to
6 implementing AfL in their classes. The results indicated that PE teachers felt there was little
7 coherence between the learning curriculum objectives they set for their pupils and the
8 assessment methods they used. In addition, Borghouts et al. (2016) reported that teachers
9 considered AfL important, but their actual assessment practices did not reflect this.

10 Also, FA has been associated with an AA approach (Dinan-Thompson & Penney, 2015; Hay
11 & Penney, 2009; Penney et al., 2009). The concept of AA refers to assessment techniques,
12 instruments and activities that are clearly applied within real situations, activities and learning
13 contexts. AA is in direct contrast to any specific and artificial assessment situations that bears
14 no relation to real practice or the actual application of appropriate knowledge. According to
15 López-Pastor et al. (2013), the main characteristics of AA in PE are as follows. First, AA refers
16 to the use of assessment techniques, instruments and activities applied to real situations or that
17 are transferable to real life outside the classroom. Second, AA and feedback processes are
18 carried out on a regular basis. Third, AA involves the pupils in the assessment of their learning
19 (self-assessment, peer assessment, co-assessment, etc.).

20 Godbout (2001), for example, indicated that carrying out AA implies having criteria that
21 differentiate it from other assessment systems. For example, AA promotes complex
22 knowledge, encourages reflection as a mental routine, records the evolution of the pupil's
23 knowledge, and uses the assessment processes as learning moments. Mintah (2003) stated that
24 AA is quite widespread among PE teachers due to the characteristics of the subject and that its
25 use helps pupils improve their self-concept, their motivation and their skill level.

1 Also, Hay and Penney (2009, p. 395) consider that AA should be based in movement and
2 capture the cognitive and psychomotor processes involved in the competent performance of
3 physical activities. In addition, these authors state that assessment should redress the
4 mind/body dualism propagated by traditional approaches to assessment, curriculum and
5 pedagogies in PE, through tasks that acknowledge and bring to the fore the interrelatedness of
6 knowledge, process (cognitive and motor), skills and the affective domain.

7

8 **Methodology**

9 ***Design and sample:***

10 The study took place in schools in a city in Spain. The Spanish government sets minimum and
11 common standards to PE, but each school and teacher has a great deal of autonomy to complete
12 and adapt them to their context. In this study we focus on a number of teachers who had taken
13 advantage of this freedom to become assessment-literate in using alternatives to the traditional
14 approach.

15 Semi-structured interviews were conducted with a purposive sample of four PE teachers
16 experienced in the use of assessment: three from Primary Education (one woman and two men)
17 and one from Secondary Education (one woman). The teachers met four criteria when assessing
18 pupils' learning, regardless of the education setting in which they taught: a) they used
19 techniques, instruments and assessment activities applied to real situations or that have
20 transferability to real life outside the classroom; b) they were capable of assessing a range of
21 skills and competences applicable to situations in everyday life; c) they facilitated FA
22 processes; and d) they involved pupils in the assessment of their learning. Furthermore, these
23 teachers were of special interest because despite of some differences between them in terms of
24 levels of knowledge about FA and AA, they all had a good knowledge in assessment given
25 their involvement in a community of practice interested in FA. On the one hand, this fact could

1 be seen as a limitation of the study, but on the other hand the physical education community
 2 can learn from such teachers and see the importance of the support a community of practice
 3 provides them.

4 In Table 1 we briefly outline the professional characteristics of each participant.

5 **Table 1.** Professional characteristics of the teachers interviewed.

	Teacher 1	Teacher 2	Teacher 3	Teacher 4
Age	36	38	43	41
Initial qualification.	Master degree in PE.	Bachelor degree in PE.	Bachelor degree in PE.	Bachelor degree in PE.
Additional qualifications.	PhD.	Master degree in PE.	-Pedagogy. -PhD Training.	-Pedagogy. -PhD (thesis on FA).
Experience of FA and AA during PETE programmes	None.	None.	None.	Yes, in some modules.
Professional development.	-Is part of a community of practice that uses FA. -Contact with others that use FA. -Has read books and articles on the subject. -Has attended conferences on assessment.	- Is part of a community of practice and colleagues that use FA. -Has read books and articles on the subject. -Has attended conferences on assessment and alternative assessment methods.	- Is part of a community of practice that use FA. -Has read books and articles on the subject.. -Has attended conferences.	-PhD Thesis on FA in Primary PE. -Has completed professional development courses. -Is part of a community of practice on FA in education. -Has attended conferences.
Teaching experience.	3 years in university and 2 years in Secondary.	4 years in Secondary and 8 in Primary.	20 years in Primary PE.	12 years in PE and other subject areas in Primary.
Time spent using FA or similar methods.	In Secondary 2 years.	8 years since starting in Primary.	15 years.	12 years.
Average of pupils per class.	16 – 24.	18 – 22.	Around 25 per class.	17-18.

6 Source: prepared by the authors

7 **Data collection methods:**

1 Semi-structured interviews with each of the teachers were carried out according to Kvale's
2 (2007) guidelines for qualitative interviews, to gather information on their assessment
3 practices. Each teacher was sent a letter inviting them to participate in the study as well as
4 written information about the study. The teachers were required to give their informed consent
5 to participate, and were assured confidentiality in relation to the management, analysis and
6 reporting of their data (using a number to refer a every teacher). All interviews were tape-
7 recorded and conducted by the first author, who also transcribed the full interviews. Once the
8 interviews were recorded, a word-for-word transcription was made. The interviews began with
9 some questions related to personal information about the interviewees. The second part related
10 to key aspects of the research aims such as the assessment system they used, how they recorded
11 the information generated by the process, and the reasons that had led them to carry out that
12 type of assessment. All interviews began by providing guarantees of anonymity and
13 confidentiality of the answers provided and interviewees were asked to give their authorization
14 for the interviews to be recorded (Stake, 2010).
15 Each interview lasted between 1 hour and 1 hour and a half. It had 26 questions which could
16 be extended if the information provided was not enough or we wanted to deepen any aspect. In
17 the table 2 we present some of the key questions.

18 **Table 2.** Key questions of the interview

1-What kind of assessment instruments do you use? Why those? 2-Do you use direct observation as the main way of recording information to assess? Or, do you use more structured assessment instruments? Why do you use one or the other? What are the advantages and constraints you think each system has? 3-Are your assessment instruments integrated within the learning and teaching processes? 4-How and when do you give feedback with the assessment instruments that you use? 5-Do you use self-assessment processes? What instruments do you use? Why do you do it? Do you use peer-assessment processes? What instruments do you use? Why do you do it? Do you use collaborative-assessment processes? What instruments do you use? Why do you do it? 6-How many pupils have you got in each group? How does group size influence the type of you use? 7-Do you do learning oriented assessment, where all your feedback is oriented to your pupils' learning? 8-If you have not experienced formative assessment and/or authentic assessment in your pre-service Teacher Education, why do you develop it as a teacher? How did you come to this type of assessment practice (FA and AA)? What type of In-Service Teacher Education has led you to use this type of assessment (readings, courses, congresses, teacher groups, seminars, talk with other teacher colleagues, etc.)?

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Data analysis:

Data from the interviews were analyzed and categorized according to the constant comparative method from a constructivist theory approach (Charmaz, 2006). Data collected during the interviews were summarized into different themes using open coding, which were confirmed and modified throughout the analysis. Specifically, the data analysis began just after the first interview and was carried out using open, axial and selective coding. The constant comparative analysis generated a number of themes and subthemes. For instance, in the first analysis the two themes were: 'How teachers described their successful use of and experiences of FA and AA' and 'How these teachers recorded FA and AA evidence'. In the first theme the teachers' responses were grouped into these characteristics: (1.1) Assessment applied to real situations or that are transferable to real life; (1.2) Formative assessment and frequent feedback; (1.3) Pupil participation in the assessment. And the second theme was initially based on three subthemes: (2.1) Recording observations; (2.2) Positive aspects of recording observations; (2.3) Dealing with difficult situations. This initial coding in which provisional themes were identified was useful to be applied in the interviews that followed, to verify that they were accurate and appropriate. The axial coding was done with the objective of finding connections between the identified themes. Finally, selective and theoretical coding was carried out to discover how the themes evolved from identifying similarities and differences among the themes and subthemes, until saturation was reached in the moment no new themes emerged from the data, formulating afterwards a new conceptualizations of assessment practices (Charmaz, 2006). Table 3 shows the themes and subthemes that were finally constructed from the data.

Table 3. Themes and subthemes constructed from the data.

Themes	Subthemes
1- How teachers described their challenges with the use of FA and AA.	1.1-Assessment applied to real situations or that are transferable to real life. 1.2-Formative assessment and frequent feedback. 1.3-Pupil participation in the assessment.
2- How these teachers recorded FA and AA evidence.	2.1-Recording observations. 2.2-Positive aspects of recording observations. 2.3-Dealing with difficult situations.

1

2 Despite a constructivist approach emphasized the construction of theories in the process of
3 interaction between the researcher and the informant people, this research try also to give
4 credibility to the research the data were analysed by two researchers independently, one of
5 whom was an experienced qualitative researcher, and then discussed and reworked until
6 agreement was reached over themes and subthemes. The final interpretations were revised by
7 the interviewees through respondent validation to confirm their veracity. Regarding
8 applicability or transferability, a large number of descriptive data of the contexts in which the
9 teachers worked was also generated. A further step in the process of confirming the reliability
10 of the data analysis was the participation of an experienced qualitative researcher external to
11 the study, who reviewed the fit between themes and subthemes and the data.

12

13 **Results**

14 Results are organized in two main themes related to the research objectives. Each theme is
15 divided into subthemes indicated in the table 3.

16 ***How teachers described their challenges with the use of FA and AA***

17 *Assessment applied to real situations or that are transferable to real life*

18 The interviewees said that they used instruments and assessment tasks that are applicable to
19 real situations or that are transferable to real life outside the classroom:

1 *The assessment instruments are integrated into the teaching and learning process; for*
2 *example: recording the lap times in the activity of learning to run to a pace, checking*
3 *heart rate, if they want to run after school outside of the classroom, in real life.*
4 *Therefore, I always look for something that hooks them, or they like enough to do*
5 *outside of class and that helps improve aspects of their life, such as avoiding a*
6 *sedentary lifestyle (T1).*

7 This teacher made reference to how they used the assessment instruments in their daily classes,
8 giving formative feedback and strategies to the students to improve their running pace in the
9 unit and in real life contexts. In contrast with traditional assessment practices, they did not use
10 a physical running test to grade their pupils; they assessed mainly the entire learning process
11 throughout the unit. Thus, the final grade is based on whether or not the pupils learnt how to
12 run efficiently, at a sustainable pace, and not merely on how fast or slow they were.

13 In another case, a teacher noted a transfer of learning to their real life:

14 *The pupils carry out activities that are useful, so that they can do them in their daily*
15 *life. For example: the rubric of self-evaluation makes them reflect and learn to have*
16 *autonomy and responsibility (T2).*

17

18 *Formative Assessment and frequent feedback*

19 Two teachers said that they delivered formative and regular feedback to their pupils about the
20 activities they undertook. They offered three types of feedback to them, whole group, small
21 group, or individual feedback:

22 *The feedback can be immediate or later and, mainly, it is verbal. Unless they produce*
23 *an assignment that I correct and then I give them the feedback. Sometimes I give it in a*
24 *group (if a high percentage of the class has not understood what was required) and*
25 *sometimes individual or in small groups (if it is a specific thing) (T1).*

1 *I give feedback to the pupils in very quick individual interviews or provide immediate*
2 *corrections about key aspects of what they are doing. Also, sometimes I stop the activity,*
3 *pause to explain something and then continue (T4).*

4 The feedback described by these teachers was continuous, fully integrated into the teaching-
5 learning processes and clearly focused on pupils improving their practice and learning. This
6 contrasts with traditional assessment practices, in which feedback tends to be provided towards
7 the end of a unit with a summative and grading goal.

8 *Pupil participation in assessment*

9 Assessment was carried out in a manner in which pupil participation was encouraged, whilst
10 collecting the maximum amount of data about the learning process was a priority. Pupils often
11 used self-assessment and peer-assessment instruments during the daily learning processes:

12 *When I want to do peer-assessment with the children, I think a descriptive scale works*
13 *very well, especially beginning with what is the assessment process. If you have the*
14 *items marked and each of them has its subscales (a, b, c), you guide the pupil's*
15 *behaviour and find it easier to determine what to observe (putting a cross to indicate*
16 *their response) (T2).*

17 *The pupils have a self-assessment questionnaire that they use in each teaching unit*
18 *(sometimes at the end, or in the middle and at the end, and sometimes at the beginning,*
19 *in the middle and at the end), the good thing about providing it at the beginning it is*
20 *that they 'connect ' a little more, they know what they have to focus on to improve in*
21 *the areas that you want them to (T3).*

22 The data shows how teachers believed that peer and self-assessment improve student learning
23 and may be used effectively by pupils. But, as in other procedures, pupils need to be taught to
24 do this. This learning process can be facilitated by providing students with assessment
25 indicators, using them regularly, offering different experiences and using the same instrument

1 at different times in the same teaching unit (the first day, in the middle and at the end of the
2 unit).

3 In other cases, self-assessment was also used as a final assessment process of the term: the
4 pupils were informed of the criteria from the beginning of the school year.

5 *I use self-assessment rubrics at the end of each term, the descriptive scale type, with*
6 *very clear items and with several levels to choose from. This instrument is given to the*
7 *pupils at the beginning of the term, so they know what the most important aspects of*
8 *learning are (T2).*

9 As such, pupils were aware from the beginning of the term of the bases on which their learning
10 was being judged, both by themselves, their peers and their teacher. Among other things, this
11 led to enhanced engagement in lessons by most pupils.

12 At the end of each term, the teachers carried out a dialogue with their pupils, in the form of
13 informal interviews that served to adjust the teachers' views on learning. All four teachers felt
14 FA and AA supported the progressive improvement of student learning:

15 *Yes, in the self-grading questionnaire that they do at the end of the term, there was a*
16 *big difference between what some pupils and I thought. We did interviews with these*
17 *pupils (if the group is large), or with everyone (if the group is smaller). These have*
18 *been the most fruitful years, where I have noticed a lot more change in them, even in*
19 *just 10 minutes at break time ... and there were changes in their attitudes towards me*
20 *(T3).*

21 In addition, at the end of each term, pupils undertook a written and reasoned self-grading
22 process. Subsequently, interviews are conducted using 'dialogic grading', in which the final
23 grading was discussed and agreed between the pupil and the teacher:

24 *They are assessed quarterly and then we conduct personal interviews. The assessment*
25 *instrument indicators are the same as those that are used later to grade them, there is*

1 usually no difference of more than 1 point. When the pupils self-grade and there is more
2 than 1 point of difference between their perception and mine, we go to a co-assessment
3 interview; they know that their opinion about their work counts towards their grade
4 (T1).

5 In conclusion to this section we can say that in these teachers' experiences, learning objectives
6 and the assessment criteria allow pupils to be involved as co-constructors of their learning.
7 Objectives and criteria also help pupils to carry out reliable self-assessment processes in real
8 situations or transferable to real life. Moreover, these teachers only used grades at the end of
9 the process (unit or term) with self-assessment and self-grading techniques that were regulated
10 by dialogic processes between teacher and pupils. In all these cases, teachers used FA during
11 the learning process, and only 'grade' at the end, when the administration requires a grade for
12 every student in each of the subjects.

13

14 ***How these teachers recorded evidence through FA and AA***

15 Also, within this second theme we can check how this group of teachers said that they used
16 three features of assessment: recording information, making judgements and reporting
17 information to stakeholders.

18

19 *Recording observations*

20 Teachers explained that they used direct observation as a routine technique for collecting
21 information, within which they used a range of different instruments. They commented that
22 they used semi-structured or unstructured instruments such as teacher's notebook, teacher's
23 diary, rubrics (for instance, a descriptive scale with 3-4 levels), etc.:

1 *I use direct observation, it is inevitable; which is recorded in the teacher's notebook,*
2 *except for the days that have a rubric to record the information, when this instrument*
3 *is used (T1).*

4 In addition, two teachers said that they used other more structured assessment tools to record
5 information from direct observation. They reflected on the key aspects to be assessed from
6 each teaching unit. For example: rubrics, individual and group tracking cards, verbal or
7 numerical scales, checklists, etc.:

8 *I use tracking cards, I observe the pupils based on them, unless something*
9 *extraordinary happens that I also pick up on (T3.)*

10 *In the observation sheets (for group monitoring) I used some very basic observations*
11 *to indicate if the items and objectives were being met and the degree of compliance:*
12 *very basic scale: very good, good, fair, bad. And if the goal that we set was achieved*
13 *or not (T4).*

14 In addition, teachers were beginning to use Information and Communication Technologies
15 (ICT) in their FA and AA processes. For example, some replaced their traditional paper-based
16 teacher's notebook with ICT tools, such as tablets:

17 *A teacher's notebook (it's a file inside the Tablet), with a section of observations where*
18 *I usually write something (T2).*

19 *By using the tablet in class, I record a lot and so I'm able to see it again. For example,*
20 *when skipping, I recorded them in 'slow-motion', which is useful for giving them*
21 *corrective feedback, since they can see where they made mistakes and how they should*
22 *do it the next time. This example is only visual feedback for the pupil using video*
23 *recordings, but it is not recorded in any paper form (T2).*

24 The use of ICT to provide feedback to students in the moment was a very useful practice to
25 support learning. Unlike traditional assessment systems in which the process is little taken into

1 account, ICTs can facilitate the registration of aspects we want to assess and provide other
2 feedback, such as visual information, immediately.

3 The examples in this section show a different way of recording information from observation
4 regarding traditional forms of assessment. If ICT is used to do traditional assessment, it is
5 usually for the same purpose, without seeking feedback that helps inform learning.

6

7 *Positive aspects of recording observations*

8 Teachers made positive comments about being able to record observed behaviours. In some
9 cases, they helped pupils to be aware of their own learning process, therefore becoming more
10 directly involved in it:

11 *The rubric is clearer for the pupil; they know perfectly what you are going to assess*
12 *them on, what you will ask of them. I believe that it provides focus on the learning*
13 *objective (T1).*

14 In other cases, it was the teachers themselves who indicated that the rubric helped them retain
15 some appreciation of what really worked in class and whether the process was positive or not:

16 *Recording observations based on the criteria established in some tracking sheets*
17 *helped me a lot. The pupils saw that we were working on concrete aspects of PE and*
18 *that we were not in the playground just playing whatever we wanted, in short, it helps*
19 *me focus and get them to focus on what's important (T4).*

20 An essential part of carrying out co-assessment with pupils was that they were aware of and
21 participated in the creation of the learning objectives that they aimed to achieve. But assessment
22 had a powerful influence on how pupils learn in PE, as this teacher noted:

23 *For example: if I pose two or three challenges and have only got time to do one or two,*
24 *I record it on the Tablet, as a note for the teacher, and the following year I'll pose fewer*
25 *challenges. The important thing is that this information will serve you for other years,*

1 *so as not to make the same mistakes and you can take into account the difference in the*
2 *performance of the same activities by different groups of the same level (T2).*

3 This comment shows that FA is also useful to the teacher to evaluate their own practice in order
4 to improve it progressively.

5
6 *Dealing with difficult situations to record the information*

7 In general, the teachers in our study did not identify any disadvantages of recording direct
8 observation as part of an FA and AA system. However, one possible situation in daily practice
9 was the difficulty to record all the learning issues observed. Most teachers mentioned the
10 number of pupils per class as one such challenge. The time taken to fill out these assessment
11 forms, especially during classes, was also considered a problem.

12 *It is true that when you only have a few pupils you can afford the luxury of using all the*
13 *most comprehensive forms, you have to find an intermediate point (T2).*

14 The data collected shows two main drawbacks for FA and AA. One was the lack of time during
15 and/or after classes to record the information on the key aspects of learning. This was especially
16 complex when the sessions were short (45-50 minutes, including moving rooms and changing
17 into PE kit) and followed one after another without time for the teacher to write anything down.
18 A related issue was the more time that passed before they could write down the information.

19 *Yes, if you really want you can do it with 25 pupils, but when the ratio is smaller I think*
20 *it can be better.... When you have fewer pupils you work better, you can do better for*
21 *the pupils, give them more attention (T4).*

22 In spite of this, the teachers sought solutions and alternatives, such as writing down the most
23 important information on 4-5 pupils per session, using integrated assessment instruments
24 within the process completed by the pupils, and even using the time moving between
25 classrooms to carry out personal interviews with the pupils that needed feedback.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Discussion

The main contribution to the first objective for this study is to show what challenges these teachers described with the use of FA and AA in their regular PE classes. We note that these teachers attempts to meet these challenges contrast with other studies which show the constraints and failures that PE teachers experience when using this type of assessment (Dinan-Thompson & Penney, 2015; Mintah, 2003; Slingerland et al., 2017), or studies that analyse the difficulty to change the traditional approaches to assessment in PE (Lorente-Catalán & Kirk, 2016; Dinan-Thompson & Penney, 2015). In the case of the teachers in this study, participation in continuous professional development and reflexive practice in action-research groups seems important to implement FA and AA (Slingerland et al., 2017).

According to the second research objective, the study teachers affirmed that they used more systematic assessment instruments than simple unstructured observation. They used a variety of FA techniques and instruments that allowed them to collect and record data from direct observation, for which they used a range of different instruments, semi-structured or structured instruments, such as a teacher's notebook, teacher's diary, rubrics, etc. Similar results can be found in other studies (Lorente-Catalán & Kirk, 2014a; MacPhail & Halbert, 2010; Desrosiers et al., 1997). These data contradict some studies that suggest PE teachers usually use direct observation as a means to assess their pupils, but they do not usually record that information using assessment instruments, due to the time it takes to do so and the large number of pupils per class (Mintah, 2003).

In addition, it can be observed that teachers were beginning to use ICT in their FA and AA processes. These new tools allowed different FA techniques to be employed. They could also be used to write and record observations and reflections, make video recordings, take photos, apply APP, use tutorial videos to feedback and for pupils self-regulation lessons. Some

1 examples of how to use ICT can be found in Kok, Komen, Van Capelleveen and Van der Kamp
2 (2020). In fact, it is reasonable that the use of ICT become more frequent in PE given that
3 currently teachers are increasingly familiar with its use. Teachers made positive comments
4 about being able to record observed behaviours. In some cases, they helped pupils to be aware
5 of their own learning process, therefore becoming more directly involved in it.

6 An essential part of carrying out co-assessment with pupils was that they were aware of and
7 participated in the creation of the learning objectives that they aimed to achieve. When pupils
8 are informed of the assessment criteria at the beginning of each new learning process (teaching
9 unit or term), this meant that the feedback they received was related to the key learning
10 objectives, that direct observation was recorded in an assessment instrument (rubric,
11 monitoring sheet, assessment wheel) and teachers believed that the pupils took the PE activities
12 more seriously. This finding matches with the proposed aspects in MacPhail and Halbert
13 (2010), Ní Chróinín and Cosgrave (2013). As well, Redelius and Hay (2009) considered that
14 assessment had a powerful influence on pupils' approaches to and understanding of their own
15 learning in PE.

16 Teachers in this study carried out structured assessment and organized learning, obtaining
17 positive results for teachers and pupils. In contrast, Van der Mars et al. (2018) concluded that
18 the four secondary teachers observed, employed mostly informal assessment strategies and
19 graded student efforts on managerial aspects of performance.

20 Teachers improved their teaching process, but at the beginning they invested a lot of time
21 creating the assessment instruments, writing down the observations and giving feedback during
22 the process. Finally, they spent less time in the long term in the planning of activities and
23 sessions and clarifying the assessment criteria that their pupils had to attain. According to the
24 teachers, the pupils were able to participate in their own assessment and take greater interest in
25 their classes. These findings are consistent with research reported by MacPhail and Halbert

1 (2010) and Ní Chróinín and Cosgrave (2013). These teachers identified some challenges
2 recording information. For instance, in the cases of a big number of students per group, these
3 teachers registered less direct observations from each pupil. But also they identified advantages
4 like the possibility pupils were able to receive feedback from self and peer-assessment. Similar
5 situations have been reported in other studies (López-Pastor et al., 2013; MacPhail & Halbert,
6 2010; Mintah, 2003; Slingerland et al., 2017; Dinan-Thompson & Penney, 2015).

7 In spite of this, the teachers in this study sought solutions and alternatives: writing down the
8 most important information on 4-5 pupils per session, using integrated assessment instruments
9 within the process completed by the pupils, using the time moving between classrooms to carry
10 out personal interviews with the pupils that need them, record the observations as soon as
11 possible after each class, etc. Some similar ideas can be found in other studies (Desrosiers et
12 al., 1997; López-Pastor, 2006a). These authors emphasize the use of self-assessment and peer-
13 assessment during the unit as a solution to this type of issue (e. g. large number of classes, size
14 of the groups, limitations in the PE delivery hours); although, on the other hand, there are some
15 authors who point out the difficulties that appeared when pupils are involved in their own
16 assessment (Dinan-Thompson & Penney, 2015).

17 Next, we will summarise the lessons learned from these four teachers' experience with the
18 objective to know what aspects can assist or hinder the application of FA and AA in PE.

19

20 ***Key learning points***

21 *The importance of using authentic and assessment tasks integrated in the daily routine*

22 All teachers agreed about the importance of using instruments and assessment tasks that were
23 applicable to real situations or that were transferable to real life outside the classroom. In most
24 of the cases, these instruments were integrated in the daily routine of the classes and aligned
25 with curriculum requirements. The teachers insisted that FA and AA had to be integrated in a

1 curricular alignment in the learning process, and not being specific activities, different and
2 separate from the learning activities. There were many examples throughout the data, for
3 instance the teacher statement that made reference to a learning unit whose goal was that
4 students learnt how to run efficiently and with self-regulation with a sustained aerobic pace, in
5 order to apply this new skill to daily life. This practical proposal for efficient running was stated
6 by the professional commitment of the teachers inspired by the Barba and López-Pastor
7 (2006b) pedagogical proposal, not as a government requirement; it is not a learning outcome
8 stipulated for the subject. In the first and second sections of the findings, we explained how the
9 teachers used integrated assessment within the teaching and learning process. These results
10 matched what Godbout (2001) found on the relationship between integrated assessment and
11 AA, when he explained the four criteria that differentiate AA from other assessment systems,
12 the final one of which referred to the use of assessment as learning moments. Also, López-
13 Pastor et al. (2013) point out that AA is an assessment system closely related to FA, in which
14 learning activities must be the same or in similar conditions to those activities developed in
15 real life, so the assessment procedures and instruments should be capable of identifying the
16 different skills and competencies that are used in these situations. In addition, delivering
17 authentic and useful assessment is a pre-condition of ensuring that assessment is effective (Hay
18 and Penney, 2009). These results also match the findings of other studies (Ní Chróinín &
19 Cosgrave, 2013; MacPhail & Murphy, 2017) indicating that carrying out assessment integrated
20 within the teaching and learning processes generates improved learning in the pupils.

21 *Formative and regular feedback as fundamental to learn more effectively*

22 The interviewees discuss that they delivered formative and regular feedback to their pupils
23 about the activities they undertook. The use of regular feedback was one of the fundamental
24 premises for carrying out FA, and some authors even insist that formative feedback is the key
25 to differentiating a true FA process from one that is not (Black & Wiliam, 2009; Dinan-

1 Thompson & Penney, 2015; Slingerland et al., 2017). Moreover, the teachers in this study
2 believed this approach to providing feedback helps pupils to learn more effectively, involving
3 them as con-constructors of their own learning (Leirhaug & MacPhail, 2015; Tolgfors &
4 Öhman, 2015).

5 *It is important to encourage students' participation in assessment*

6 Teachers highlighted the importance of encouraging pupil participation in assessment. Pupils
7 often used self-assessment and peer-assessment instruments during the daily learning
8 processes. Other studies also promoted different types of pupil involvement in the assessment
9 processes (self-assessment, peer-assessment and co-assessment) (Desrosiers et al, 1997;
10 López-Pastor, 2006a; Herranz & López-Pastor, 2014b; Lorente-Catalán & Kirk, 2016). In this
11 sense, the teachers' views of their assessment practices in this study were in line with a wide
12 range of evidence accumulated over the last 20 years about how this type of assessment
13 generates improvements in the engagement of primary and secondary pupils (Herranz &
14 López-Pastor, 2014b; Leirhaug & MacPhail, 2015; López-Pastor, 2006a; MacPhail & Halbert,
15 2010; Ní Chróinín & Cosgrave, 2013; Tolgfors & Öhman, 2015). The teachers in our study
16 match with Black and Wiliam (1998) in the use of self-assessment as one of the basic elements
17 that enable pupils to be aware of their level of learning and participate in their own assessment.
18 However, teachers in this study seem to have resolved, to some extent, the issues that Black
19 and Wiliam (1998) indicated; in this sense, they have the knowledge about how to use pupil
20 self-assessment and avoided the distrust generated when pupils can participate actively in their
21 own assessment. Furthermore, the results seem consistent with the later approaches of Black
22 and Wiliam (2009), concerning the issue that self-assessment should not be restricted to a
23 specific moment within the assessment process as this would detract from its formative
24 potential. It is therefore preferable to ensure self-assessment is ongoing throughout the lesson,
25 so that pupils can review their work continuously and not only at the end. In this study, self-

1 assessment with formative and summative purposes co-existed in a complementary way, as we
2 noted on pages 17-18.

3 Sometimes, the teachers in this study also affirmed that they used self-assessment instruments
4 at the end of the term. In these cases, pupils were informed about the assessment criteria from
5 the beginning of the school year. When teachers used a written and reasoned self-grading
6 process, subsequently interviews were conducted using 'dialogic grading', in which the final
7 grading was discussed and agreed between the pupil and the teacher. These practices contrast
8 with Black and Wiliam (2009) affirmations, who consider self-grading is not an essential
9 element and that it may even be counterproductive in the teaching and learning processes, since
10 pupils stop focusing on the really important aspects of learning and give priority to the mark.
11 This did not seem to be the case with the teachers in this study, who consistent with other
12 studies (Herranz & López-Pastor, 2014b), thought that if pupils were given the responsibility
13 of self-grading, it usually brought benefits for them through taking greater responsibility for
14 their own learning and improving the way they performed tasks.

15

16 **Conclusion**

17 We began this paper saying that previous research has indicated that assessment has proven to
18 be problematic in PE teaching and FA, in particular, seemed not being implemented in the
19 classroom. For this reason, the first research objective of this study was to identify the
20 challenges with using FA and AA in PE teaching, based on the perspectives of some teachers
21 experienced in using assessment in their classes. After listening to and analyzing the words of
22 these teachers, we consider that we have shown how these teachers face their challenges with
23 using FA and AA: how they give formative and frequent feedback, encourage the participation
24 of pupils in the assessment processes, apply the assessment in real or transferable situations to
25 real life outside the classroom and share with students the learning objectives and assessment

1 criteria. All this helps pupils to become more involved in their learning process in the subject
2 and to be able to carry out fairly reliable self-assessment processes. The second research
3 objective was to analyse how these teachers recorded evidence through FA and AA. The
4 teachers in this study encountered some challenges with recording information, but also they
5 provided some solutions and alternatives when developing FA and AA in PE. Nevertheless,
6 the teachers also emphasised that to obtain positive results for teachers and pupils, it is key that
7 the teaching is organised with assessment in mind, and that the assessment is structured. For
8 example, with the use of rubrics, pupils do self-assessment or peer-assessment which stimulates
9 two things: (a) a better focus on what they have to learn; (b) improvement of the self-regulation
10 of learning.

11 This study can be useful to PE teachers who want to apply FA and AA within their classes,
12 knowing some strategies to resolve challenges that may arise from it. Teachers consider that
13 they obtain two great educational and curriculum alignment advantages working like this. One
14 is that FA and AA improves the pupils' learning in PE. The second is pupils are more aware of
15 the learning objectives they have to achieve and take their learning more seriously.
16 Furthermore, the teachers improve their teaching process, using less time in the medium-long
17 term doing their planning and establishing clearer assessment criteria for the pupils to achieve.
18 However, they indicate some constraints when collecting the observation data, and offer some
19 strategies that help to overcome them (e.g. promoting the autonomy and participation of the
20 pupil in their own assessment with strategies, such as peer-assessment and self-assessment).
21 Many of these difficulties are pointed out in the international literature of PE as drawbacks that
22 prevent the use of FA and AA to many PE teachers. What this study seems to show is that with
23 pedagogic knowledge and positive attitude to use FA and AA it is possible to overcome these

1 drawbacks. But more research observation based on showing how teachers apply FA and AA
2 into the classroom is needed.

3 Although what we learnt from these teachers is encouraging, it should be taken into account
4 that they had previous knowledge and experience about FA and AA, possibly more than an
5 average PE teacher. They belonged to communities of practice linked to assessment practices.
6 This reinforces the idea of the need, defended by numerous authors and recently by AIESEP
7 (2020), to educate teachers specifically in this subject, either from Initial Teacher Education
8 and above all, from continuous professional development. But also, there is a need for teachers
9 to be supported by communities of practices that accompany and educate teachers in the
10 process of change. This seems to be a factor to take into account to reach the desired
11 “assessment literacy”. The challenge now is to find ways to involve teachers in communities
12 of practice that help them to initiate those changes that should lead, in turn, to quality PE.

13 **References**

- 14 AIESEP, (2020). *Position Statement on Physical Education Assessment*. Retrieved from:
15 <https://aiesep.org/scientific-meetings/position-statements/>
- 16 Barba, J.J., & López-Pastor, V.M. (2006). *Aprender a correr con Autonomía. Unidades*
17 *didácticas en primaria y secundaria (Learning how to run autonomously. Didactic units*
18 *in Primary and Secondary school)*. Buenos Aires: Miño y Dávila.
- 19 Bernstein, B. 1977. *Class codes and control, towards a theory of educational transmissions,*
20 *Vol. 3*. London: Routledge and Keegan Paul.
- 21 Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education:*
22 *principles, policy & practice*, 5(1), 7-74.
- 23 Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational*
24 *Assessment Evaluation and Accountability*, 21(1), 5-31.

- 1 Borghouts, L., Slingerland, M., & Haerens, L. (2016). Assessment quality and practices in
2 secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, 1-17. doi:
3 10.1080/17408989.2016.1241226.
- 4 Casbon, C., & Spackman, L.S. (2005). *Assessment for learning in Physical Education*. Leeds:
5 BAALPE.
- 6 Charmaz, K. (2006). *Constructing Grounded Theory*. London: Sage.
- 7 Chng, L., & Lund, J. (2018). Assessment for Learning in Physical Education: The What, Why
8 and How. *Journal of Physical Education, Recreation and Dance*, 89(8) 29-34
- 9 Desrosiers, P., Godbout, P., & Genet-Volet, I. (1997) Teachers' Assessment Practices Viewed
10 Through the Instruments Used in Physical Education Classes. *Journal of Teaching in*
11 *Physical Education*, 16(2), 11-228.
- 12 Dinan-Thompson, M., & Penney, D. (2015) Assessment literacy in primary physical education.
13 *European Physical Education Review*, 21(4), 485-503.
- 14 Godbout, P. (2001). Reflection on Innovation and Renovation: Authentic Assessment and
15 Subject Matter Expertise in Physical Education. *AIESEP Congress*, Madeira.
16 [HTTP://aiesep.org/wp-content/uploads/2014/11/Paul-Godbout.pdf](http://aiesep.org/wp-content/uploads/2014/11/Paul-Godbout.pdf)
- 17 Hay, P. (2006). Assessment for learning in physical education. In: Kirk, D., Macdonald, D. &
18 O'Sullivan, M. (ed). *The handbook of physical education*. London: Sage, pp. 312–325.
- 19 Hay, P., & Penney, D. (2009). Proposing conditions for assessment efficacy in Physical
20 Education. *European Physical Education Review*, 15(3), 389-405.
- 21 Hay, P., & Penney, D. (2012). *Assessment in physical education: A sociocultural perspective*.
22 Routledge.
- 23 Herranz, M., & López-Pastor, V.M. (2014b) ¿Es Viable Llevar a Cabo Procesos de
24 Autoevaluación y Evaluación Compartida en Educación Física en la Etapa de
25 Educación Primaria? Un Estudio de Caso Longitudinal (Is it Feasible to Carry Out Self-

- 1 Assessment and Shared Assessment in Physical Education in Primary Education? a
2 Longitudinal Case Study). *Revista de Educación Física, Renovar la teoría y la práctica*,
3 133(30-1),11-18.
- 4 Kvale, S. (2007). *Doing interviews*. Thousand Oaks, California: Sage Publications.
- 5 Kok, M., Komen, A., Van Capelleveen, L., & Van der Kamp, J. (2020). The effects of self-
6 controlled video feedback on motor learning and self-efficacy in a Physical Education
7 setting: an exploratory study on the shot-put. *Physical Education and Sport Pedagogy*,
8 25(1), 49-66, DOI: 10.1080/17408989.2019.1688773
- 9 Leirhaug, P.E., & Annerstedt, C. (2016). Assessing with new eyes? Assessment for learning in
10 Norwegian physical education. *Physical Education and Sport Pedagogy*, 21(6), 616-
11 631.
- 12 Leirhaug, P.E.. & MacPhail, A. (2015) It's the other assessment that is the key: three Norwegian
13 physical education teachers' engagement (or not) with assessment for learning. *Sport*,
14 *Education and Society*, 20(5), 624-640.
- 15 López-Pastor, V.M. (2006). *La Evaluación en Educación Física: Revisión de los modelos*
16 *tradicionales y planteamiento de una alternativa: La evaluación formativa y*
17 *compartida. (Assessment in Physical Education: Review of traditional approach and an*
18 *alternative: formative and shared assessment)*. Buenos Aires: Miño y Dávila.
- 19 López-Pastor, V.M., Kirk, D., Lorente-Catalán, E., Mc Phail, A. & Mc Donald, D. (2013).
20 Alternative Assessment in Physical Education: A Review of International Literature.
21 *Sport, Education & Society*, 18(1), 57-76. DOI:10.1080/13573322.2012.713860.
- 22 Lorente-Catalán, E., & Kirk, D. (2014). Making the case for democratic assessment practices
23 within a critical pedagogy of physical education teacher education. *European Physical*
24 *Education Review*, 20(1), 104-119.

- 1 Lorente-Catalán, E., & Kirk, D. (2016). Student teachers' understanding and application of
2 assessment for learning during a physical education teacher education course. *European*
3 *Physical Education Review*, 22(1), 65-81.
- 4 MacPhail, A., & Halbert, J. (2010). We had to do intelligent thinking during recent PE:
5 Students' and teachers' experiences of assessment for learning in post-primary physical
6 education. *Assessment in Education: Principles, Policy & Practice*, 17(1). 23–39.
- 7 MacPhail, A., & Murphy, F. (2017). Too much freedom and autonomy in the enactment of
8 assessment? Assessment in physical education in Ireland. *Irish Educational Studies*, 1-
9 16.
- 10 Mintah, J.K. (2003). Authentic Assessment in Physical Education: Prevalence of Use and
11 Perceived Impact on Students' Self-Concept, Motivation, and Skill Achievement.
12 *Measurement in physical education and exercise science*, 7(3), 161-174.
- 13 Ní Chróinín, D., & Cosgrave, C. (2013). Implementing formative assessment in primary
14 physical education: teacher perspectives and experiences. *Physical Education and Sport*
15 *Pedagogy*, 18(2), 219-233. doi:10.1080/17408989.2012.666787.
- 16 Nyberg, G., & Larsson, H. (2014). Exploring 'what'to learn in physical education. *Physical*
17 *education and sport pedagogy*, 19(2), 123-135.
- 18 Oslin, J. L., Mitchell, S.A., & Griffin, L.L. (1998). The game performance assessment
19 instrument (GAPI): development and preliminary validation. *Journal of Teaching in*
20 *Physical Education*, 17(2), 231-243.
- 21 Penney, D., Broker, R., Hay, P. & Gillespie, L. (2009). Curriculum, pedagogy and assessment:
22 three message systems of schooling and dimensions of quality physical education.
23 *Sport, Education and Society*, 14 (4), 421 442.

- 1 Redelius, K., & Hay, P. (2009). Defining, acquiring and transacting cultural capital through
2 assessment in physical education. *European Physical Education Review*, 15(3), 275-
3 294
- 4 Richard, J.F., & Godbout, P. (2000). Formative Assessment as an integral part of the teaching-
5 learning process. *Physical and Health Education Journal*, 66(3), 4-9.
- 6 Scanlon, D., MacPhail, A., Walsh, C., & Tannehill, D. (2022). Embedding assessment in
7 learning experiences: enacting the principles of instructional alignment in physical
8 education teacher education. *Curriculum Studies in Health and Physical Education*, 1-
9 18.
- 10 Slingerland, M., Borghouts, L., Jans, L., et al. (2017). Development and optimisation of an in-
11 service teacher training programme on motivational assessment in physical education.
12 *European Physical Education Review*, 23(1), 91-109.
- 13 Stake, R. (2010). *Qualitative research: Studying how things work*. Amsterdam: The Guildford
14 Press
- 15 Svennberg, L., Meckbach, J., & Redelius, K. (2014). Exploring PE Teachers' 'Gut Feelings':
16 An Attempt to Verbalise and Discuss Teachers' Internalised Grading Criteria.
17 *European Physical Education Review*, 20(2): 199-214.
- 18 Tolgfors, B., & Ohman, M. (2015). The Implications of Assessment for Learning in Physical
19 Education and Health. *European Physical Education Review*, 22(2), 150–16.
- 20 Van der Mars, H., Timken, G., & McNamee, J. (2018). Systematic Observation of Formal
21 Assessment of Students by Teachers (SOFAST). *Physical Educator*, 75(3), 341-373.