

Student teachers' understanding and application of Assessment for Learning during a Physical Education Teacher Education course

There is widespread consensus on the need for Assessment for Learning (AfL) in both university courses and school programs. Given the prevalence of traditional practices in school physical education where assessment is basic or non-existent we might ask whether AfL is present in Physical Education Teacher Education (PETE) courses? Where it is, we may wonder whether the student teachers' concept of AfL is consistent with the concept advocated for and developed in the literature and in policy. This paper draws on a qualitative study that was conducted on a core unit within a PETE course delivered by a university in England. Despite the fact that the students lacked a voice in framing their understanding and uses of AfL, we could say that most students appeared to have grasped the concept, not just as a theoretical framework but also in terms of applying different strategies during their practice in schools. They reported that they found this knowledge relevant and useful, they recognised the need to learn more about it, and they also had the intention to incorporate this knowledge in their future professional practice. Given these findings, we conclude this paper with a brief discussion of why traditional approaches to assessment in school physical education appear to persist.

Keywords: Assessment for Learning; Physical Education; Teacher Education

Introduction

Assessment is a particularly troublesome process for Physical Education Teacher Education (PETE) given the practical nature of the subject in the compulsory years of primary and secondary school and the difficulties physical education teachers have traditionally experienced in assessing pupil learning (Hay and Penney, 2013; López-Pastor et al., 2013). Traditional assessment approaches in PE have often been product oriented, focusing on components of fitness, or de-contextualised, as in the case of assessment of isolated skills (Penney, Brooker, Hay and Gillespie, 2009). Although the wide overview of research studies on assessment in physical education provided by Lopez et al. (2013) shows important progress in this area (characterised by the authors in such terms as alternative assessment, authentic assessment, formative assessment, assessment for learning, and integrated assessment) their work also shows that these approaches are far from being regular, integral, widespread and educationally productive.

It has been argued by Biggs (1999) and Brown and Glasner (1999) among others that assessment has a central part to play in student-centred teaching. If more student-centred approaches are to be a possibility in a subject most often characterised as traditionally teacher-directed (Kirk, 2010), it is a matter of some importance to investigate whether and how student teachers learn about assessment in this subject of physical education, given that it deals predominantly with practical physical activity knowledge particularly within the compulsory years of schooling for young people between 5 and 16 years of age.

According to Penney et al. (2009), quality physical education requires an alignment of curriculum, pedagogy and assessment. **As they point out, there would be a**

two-way relationship, such that while particular assessment tasks / conditions may impact pedagogy, established pedagogical practices can also impact/ shape assessment. At the same time, it might be said that not all forms of assessment necessarily have a positive impact on pedagogy, particular assessment associated with traditional directive forms of teaching and with rote learning, which have been common practice in physical education for a century or more.

According to Hay and Penney (2009:391), for assessment to work optimally in physical education, it is necessary to integrate four conditions: a) a primary focus on assessment for learning; b) authentic assessment concerned with the relationships between learning content and contexts and their connection with the world beyond the classroom; c) assurance of validity and; d) socially just approaches to assessment. If student teachers are to learn to integrate these four conditions, they need to be present in teacher education programmes. This is not, however, a straightforward process. As revealed by Veal's (1988) research, secondary physical education teachers' actual assessment practices were rarely consistent with the practices they had been encouraged to learn during their pre-service teacher training, since the teachers believed such practices were not transferable to the reality of teaching physical education in schools, a perspective that more recent research (eg. Hay and Penney, 2009) suggests has persisted.

Beyond physical education and more generally in the university curriculum, in the 1990s and at the beginning of the 21st century some authors raised critical voices on assessment methods. They denounced assessment that was mainly in the hands of university tutors (Sluijsmans et al., 1998); students had to seek their teachers' approval regarding assessment (Fallows and Chandramahan (2001); the methods used had no consistency with academic values, goals of independent and meaningful learning, and

developing critical thinking skills (Boud, 1990); and from students' perspectives, summative assessment was prevalent in the university (Fullan, 1991). Also, other studies about dominant approaches to assessment within the university seem to indicate that few changes have been made in recent years and that this represents one of the big challenges for university teaching and learning (Zabalza, 2003). Despite López-Pastor et al.'s (2013) encouraging report, there is little reason to believe physical education teacher education is any different the situation in the university more broadly.

Boud et al., (2010) have argued that assessment is more effective when students and tutors share responsibility for it and that a more effective assessment strategy for enabling student-centred teaching is considered to be formative assessment and in particular assessment for learning (AfL), both of which imply the use of participative practices such as self-assessment, peer assessment and co-assessment. AfL is defined by the Assessment Reform Group (ARG) (2002) as the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get thereⁱ. This notion has been supported and developed by authors such as Black, Harrison, Lee, Marshall and Wiliam (2002); Blanchard (2009); Carless (2010); Sadler (2010), among others.

AfL is increasingly part of accepted orthodoxy, with massive government funding in England, a central place in national assessment in Wales, and an export to the USA (Taras, 2009). In other countries like Spain, it is a concept that has been developed especially since the last educative reform (2006) called *Ley Ordenación Educativa* (LOE). In Scotland, AfL lies at the heart of the Curriculum for Excellence (Scottish Government, 2011). Also, many different European research projectsⁱⁱ have demonstrated clear benefits accrue from the use of formative assessment, including AfL. Among these benefits can be highlighted the increased participation of students

in the learning process (Sanmartí, 2007), better results from their learning, improved processes, the development of metacognitive capacities and the aptitude for reflexive and critical analysis and, finally, the increased responsibility and autonomy of the students (López Pastor et al., 2013).

Within schools, the promotion of AfL is evident from the large number of guides and documents provided to teachers at all stages of compulsory education as part of a National Strategyⁱⁱⁱ in England. This National Strategy contains sections on assessment which include numerous sets of guidelines, resources and opportunities for sharing experiences and improving assessment, mainly based on the influential work of Black and Wiliam (1998). Although some of the ideas proposed by these authors have been called into question (Taras, 2007, 2009), they remain central to the majority of publications on the topic. Likewise, universities, in their quest for excellence, also largely define their key approach to teaching, learning and assessment by reference to the guidelines established in 1997 within the Dearing Report. These guidelines place significant emphasis on student-centred teaching methodologies and AfL. What seems clear is that both school and university policies on assessment stress that the key characteristics of AfL are as follow: a) sharing learning objectives with learners; b) helping learners to know and recognize the standards they are aiming for; c) involving learners in peer and self-assessment; d) providing feedback that leads learners to recognize their next steps and how to take them; e) promoting confidence that every learner can improve and f) involving both teacher and learner in reviewing and reflecting on assessment information (Assessment Reform Group, 1999).

Nevertheless, despite the clarity of this Assessment Reform Group (ARG) account, sometimes tutors and student teachers' perceptions of AfL do not match. MacLellan (2001), in his study of tutors' and students' perceptions of AfL in a Scottish

higher education context, a BEd (Hons) Programme), showed that there was a significant difference between tutors and students. The results suggested that while staff declared a commitment to the formative purposes of assessment and maintained that the full range of learning was frequently assessed, the students' perceptions were that they engaged in practices which militated against formative and authentic assessment.

Taking into account this literature and the policies in the context of school education and the university sector in England, it might be suggested there is an emerging consensus of the need for AfL in both university courses and school programmes. Moreover, given the strangle-hold traditional practices have on school physical education where assessment is still linked especially to the use of physical fitness tests aimed at grading the students' performance (López-Pastor et al., 2013) or is non-existent (see Moynihan et al., 2006), we might ask whether AfL is present in PETE courses and, where it is, whether the student teachers' concept of AfL is consistent with the concept advocated by and developed in the literature and in policy?

This paper draws on a qualitative study that was conducted on a core unit within a Physical Education Teacher Education course delivered by a university in England. Our purpose is to address two questions, the first concerned with the extent to which the concept of AfL in the unit investigated complied both with university policy and the literature, and the second, whether students' conceptions of AfL are consistent with their lecturers' understanding, with their university's policy, and with the research literature. We begin by providing information on the setting and participants, before outlining our research methods. Then, we investigate the university's policy on AfL before exploring in detail the students' experiences of this concept in a unit of work within their PETE programme, their concepts of AfL, and how they applied these in their practicum.

Setting and participants

The study was based in a core unit within a PETE course in a university in England. The unit is called “Assessing Learning and Developing Teaching”. It is a compulsory unit delivered in Year 2 of a 4 Year BA (Hons) Physical Education course leading to a Qualified Teacher Status (QTS) award. The unit was taught by 2 tutors and was taken by 55 students who had a one hour lecture per week and one 2-hour Seminar per week (split into 2 groups) over 5 weeks in Semester 2 (January and February 2010). One of these groups (22 students) were taken for this study. The value of using this unit for this study was that it introduced the concept of AfL to students who will eventually become physical education teachers and thus, potentially, implement this approach to assessment themselves.

The Unit “Assessing Learning and Developing Teaching”

Through this unit the student teachers learn about how they must incorporate formative assessment, including AfL in their teaching. Also, the unit proposes that students develop the ability to carry out the assessment of pupils’ progress in Physical Education both at the end of unit assessments and in GCSE and AS/A level examinations^{iv}. It also provides informal opportunities for feedback to pupils and to the teacher on the effectiveness of teaching and learning. The unit introduces the concepts of monitoring, assessment, recording and reporting in the National Curriculum Physical Education in England and the role that ICT plays in these processes.

When students have completed their lectures and seminars for the unit at the University they then undertook six weeks of teaching practice in a secondary school, which sought to provide them with the opportunity to collate evidence to show they meet the Professional Standards required for Qualified Teacher Status. It is important to note that this represents the first occasion in their course when the students are faced

with the challenge of assessing pupils. The unit assessment tasks, as we will show momentarily, are totally related with the work they will have to do in the practicum.

The aims of the Assessing Learning and Developing Teaching unit are: 1) to appreciate and understand the aims of assessment and the key characteristics of good assessment; 2) to observe and record the progress of their pupils and apply the principles of AfL, providing constructive feedback; 3) become familiar with the requirements of assessment for the national curriculum and for public examinations; 4) plan, teach and evaluate lessons for whole classes for a variety of age groups and activities as specified in the National Curriculum showing appropriate progression over 5-6 weeks work; 5) develop the skills needed to contribute to a wide range of situations in school, such as assemblies and the role of teachers in extra-curricular activities; 6) carry out a range of data collection and observation tasks through planned opportunities. The unit specification makes it clear that objectives 4, 5 and 6 are to be achieved only through the school based experience. Objective 3 is to be covered in their university classes and put into practice during their school placement.

The unit assessment strategies and tasks. According to MacPhail et al., (2010), it is important, especially in Teacher Education, to make a clear commitment to align discourse and action, or more substantively, learning, instruction, and curriculum (see also Fullan, 1991). Thus, in this case, it was necessary to articulate a strong coherence between the assessment strategies and tasks used in the unit and the AfL characteristics to give a clear message of good assessment to the students. The unit assessment comprised three assessment tasks of which only the first two affect the final unit mark. The third element, known as the Practicum, is marked only as Pass or Fail.

A) *Assessment Task 1.* Students have to design three assessment tasks that relate to the objectives of a programme element which they have designed earlier as part of the Teaching and Learning 1 unit. Students must complete and submit this task before starting their Practicum. However, they do not receive feedback on it until they have completed their Practicum. The feedback is given prior to them undertaking activity 2, which helps them improve the quality of their work.

B) *Assessment Task 2.* This task is undertaken after the students have completed their school placement Practicum. The students choose six pupils who worked through the unit they had produced for Assessment Task 1 and which was adapted to the specific needs of the pupils and the context of the school. They then critically evaluate the assessment strategy they created and implemented with their class group.

The students are notified in writing if they need to resubmit any of their work for Assessment Tasks 1 and 2 and are given feedback on the work to help them improve it.

For both Assessment Tasks 1 and 2 the tutors write their feedback on the work submitted and also complete an assessment feedback sheet which is handed to the student after they have made amendments to their Task. The feedback sheet contains a box for the tutors' final feedback to the student and suggestions for the key areas they need to develop in future. If a student fails either of these Assessment Tasks they may resubmit the work once they have corrected the errors in their initial submission.

C) *Assessment Task 3.* During their Practicum the students must apply the assessment activity they designed for Task 1 to the six pupils they chose and for whom the activity was designed. Once they have done this and completed their Practicum they must complete Assessment Task 2. However, no specific mark is awarded to this task but it becomes an additional piece of evidence in their portfolio. During their six week placement the students should also help experienced teachers implement their

assessment so they can observe how they do it. Furthermore they should observe the school teacher giving feedback in their classes and take notes on how they do it and how they question their pupils. They must write up their observations weekly and collate them for their portfolio. This evidence has to relate as closely as possible to the standards required for Qualified Teacher Status. The Practicum is assessed overall and awarded a mark of Pass or Fail. If a student fails the Practicum they can retake this element but cannot progress to the next level until they have passed it.

Method

This research was developed in two phases. It began in the academic year 2009-2010 when the first author, funded by the Spanish Ministry of Education, went to a university in England to develop a research project about assessment, pedagogy and competences in a PETE course. She had a good background on lecturing and researching on the topic of the project. The second author was the supervisor of this project with long experience on research on teacher education, assessment and curriculum studies. During the first year, a literature review on formative assessment and student participation in assessment was carried out, and the policy guidelines and programmes of the university related to assessment and specifically those related with assessment for learning were analyzed. Once we identified which lecturers and units using AfL could be appropriate for the study, the first author carried out observations of classes, attending as a participant observer in all lessons of the chosen unit (five one hour lectures and five two hour seminars). Due to the possibility to extend the research stay of the first author for one further year, students' and tutors' interviews were conducted during the second academic year (2010-2011) once the students had gone to the School Practicum.

This paper thus reports findings from a larger study whose purposes were to provide insights into how students constructed the concept of AfL and its relationship with assessment practices in physical education. In this paper, we focus on the specific research questions already mentioned and the resulting data produced by the fieldwork.

Data collection methods and analysis of data

In qualitative research there is an interconnection between data collection and analysis of data (Goetz and LeCompte, 1988; Stake, 1988). Meanwhile the researcher is observing and taking field notes he/she is already doing a sequential selection of information based on the themes that emerge during the action related to the research question. Thus, the progressing theoretical construction shapes ongoing data collection.

The methods used for data collection were:

- Document analysis. The university guidelines on teaching and learning were analyzed. Specific reference was made to the University Strategic Plan (2008-2013) and the Guides for tutors produced by the University Directorate of Teaching and Learning, particularly Guide 1 (*An introduction to Teaching, Learning and Assessment*) and Guide 3 (*A guide to Assessment for Learning*), where issues linked to formative assessment were identified. The programme and teaching materials for the unit chosen for observation were also analyzed with the aim of identifying their key design features relating to assessment.
- Observation. The first author made field notes in all the lectures and seminars attended by one of the student groups taking the unit. Observations were structured by reference to the unit objectives, the key characteristics of AfL defined by the Assessment Reform Group (1999) and the parameters defined by

the University Strategic Plan so that evidence could be obtained on the extent to which AfL is integral to the teaching of the unit.

- Interviews. Semi-structured interviews with the two tutors of the unit were undertaken according to Kvale (1996)'s guidelines to grasp how they understood the concept of AfL, if they intentionally applied AfL in the unit, and their opinions about the students' application of these strategies on the Practicum. Moreover, there was an open call to all the students who experienced the unit to be interviewed by the first author. Finally, a total of 12 of 22 students (9 female and 3 male) were interviewed using semi-structured interviews. The interviewer asked for permission from each student to record the interviews. Students' interviews were intended to discover their views on the issue and on the value of the knowledge they gained which they could apply to their teaching practice. These interviews took place the semester following when the unit was taught, after the students had completed their school placements, allowing time for them to reflect on and appreciate the impact the knowledge had.

The approach used to analyse data was a thematic analysis as it is understood by Braun and Clarke (2006). As they note, thematic analysis involves searching across a data set – be that a number of interviews or focus groups, or a range of texts – to find repeated patterns of meaning. According to these authors, a theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set (Braun and Clarke, 2006, p.10). But is also important to highlight that the thematic analysis started before the interviews in the thinking done to decide what questions to ask. These questions were validated by two physical education teacher educators who were experts in the area of evaluation.

A thematic analysis of the interviews was undertaken, paying special attention to some aspects of the data and following different phases: a) familiarising with the data, including transcription of the interviews; b) generating initial codes; c) searching for themes; d) reviewing themes; e) defining and naming themes and f) producing the written interpretation of the data.^v The field notes were analysed using a constant comparative method (Glaser and Strauss, 1967) with the parameters established by the stated unit objectives and the characteristics highlighted as being fundamental to AfL^{vi}. Additionally, the interviews were analysed using a constant comparison with the literature on AfL to identify tutors' and students' understanding of the concept, and an inductive analysis (Braun and Clarke, 2006) was carried out looking for repeated patterns of meaning, separating students' answers into different themes related to other aspects such as usefulness of learning and future intentions to use these strategies in their professional field.

Following Lincoln and Guba' (1985) recommendations, the investigators adopted a range of strategies to ensure trustworthiness. For credibility, the prolonged engagement between the researcher-observer and the participants (two years in the field) was necessary to understand the cultures she was studying. However this can also be a limitation because she may become so immersed in the culture that her professional judgments might be influenced. Also for credibility, we used triangulation of methods (observation, individual interviews and documentary analysis) and perspectives (tutors, students and observer). For transferability, we provide a complete description of the contextual factors related on the inquiry. For dependability, the processes within the study have been reported in as much detail as word limits allow. And finally, we ensured confirmability by using triangulation to reduce the effect of investigator bias and recognise our own predisposition to the research topic.

Results and discussion

We consider first of all the extent to which AfL was consistent between university teachers' understanding, university policy and the unit. We then consider the teaching of AfL within the unit, followed by the students' understanding and application of the concept.

University policy on AfL

First of all we were interested in the extent to which the concept of AfL in the unit complied both with university policy and the literature on AfL. The University Strategic Plan describes the learning and teaching strategies that tutors must employ when designing and implementing teaching and learning processes. Four key aspects were used as reference points for the observation and analysis of the target unit.

The first of these was the University's desire to align its own educational practices with those of the Bologna Process^{vii}. The second related to the need to change the current predominant teaching model, transforming it from a traditional model focused on the transmission of knowledge to one which promotes the direct and active involvement of student teachers in their learning. The third was the intention that all of the students' learning activities should be meaningful, active, reflective, collaborative and creative. The fourth and final aspect was the explicit aim of promoting innovation in assessment.

There were also 10 University Guide booklets to help tutors improve their teaching skills, plan their units and develop their programmes. Two of these were specifically designed to help with teaching, learning and assessment. Guide 1, '*An introduction to Teaching, Learning and Assessment*' (January 2009), proposed that:

‘(...) anybody involved in the teaching or assessing of our students should have a working knowledge of its contents. Whilst Guide 3 ‘*A guide to Assessment for Learning*’ (July 2009) explains: ‘(...) It is particularly aimed at those new to teaching and assessing but everyone involved in assessment should have a good working knowledge of its contents’.

This Guide details the most effective assessment practices with the intention that they should be incorporated in all subjects at all levels across the University. These guidelines were adapted by the University from ‘A Manifesto for Change’ developed by the ASKe CETL at Oxford Brookes University in 2007. According to this guide assessment has to be diagnostic, inclusive, connected with other units and balance summative and formative strategies. Also it advises that assessment tasks have to be aligned with learning outcomes, minimize plagiarism and be supported by technology. Furthermore, it explains how to engage students through applying different strategies: clarifying assessment expectations, engaging them in criteria setting; providing feedback and self-regulation through participative strategies of assessment as self-assessment and peer-assessment.

The teaching of AfL in a university unit of work

It is clearly not sufficient to consider the assessment tasks in isolation to determine whether AfL is integral to the process. It is essential to analyze what actually happens during the whole process, correlating the assessment strategy with the proposed learning activities and judging it against the accepted parameters of the concept of AfL promoted by the university and present in the literature.

To develop an interactive model^{viii} tutors should ideally discuss with students their expectations and interests and determine the dynamic of the unit through dialogue

and consensus (Blanchard, 2009). In the unit studied there was no activity in the initial class that would provide a good understanding of the students' expectations. From the tutors' point of view this wasn't necessary because they knew the students from the previous unit, but independently of that, as Blanchard suggest, doing this, tutors can gain a better understanding of their students' expectations and establish a way that enables everyone to work together. In this way students are likely to become more involved in the process of learning (Authors, 2013).

With regard to the methodology employed and its relationship to assessment, marked differences were observed between the lectures (1 hour per week) and the seminars (one class of 2 hours per week). The lectures were used primarily for the tutor to teach the content with the help of visual aids and, to a lesser extent, for the tutor to pose questions for students to reflect on individually or discuss in pairs. In lectures, the session objectives were stated at the beginning of every class and they were again reflected on at the end of each session. The University makes it clear this is a fundamental requirement. But this was only done superficially, which meant it was hard for tutors to identify if there were any real problems in students' understanding of the new concepts that had been introduced. Also, tutors had little opportunity in the lectures to provide individual feedback or verify students' learning. In contrast, the dynamic of the seminars was very different providing multiple opportunities for tutors to provide feedback to students on their work. These 2 hour sessions were used to try and apply the concepts introduced in the lectures to a series of practical tasks. The structure of each seminar was very similar. The tutor would set some questions or tasks for the students to work on in pairs or groups and then pass from group to group making comments and giving them feedback. Each pair/group then presented their conclusions to the class with the tutor ending the session by summarizing all the issues, questions

and solutions. These seminars provided the key point of contact between tutor and students, enabling the tutor to fully appreciate each student's learning and provide them with individual feedback.

Both cooperative and dialogic learning were encouraged during the unit, particularly in the seminars. In contrast, the grading was individual, based on each student's work (Assessment Tasks 1 and 2) and their Practicum. Furthermore, an interactive model would require tutors to engage in dialogue with the students to review the objectives periodically throughout the unit, and even to jointly agree the assessment criteria, which might change as a consequence of the experience, and to develop their own solutions to issues that arise. This approach, though, was not utilized by the tutors and tended to be at odds with the University's quality assurance procedures. One of the possible explanations for this could be that tutors considered it a waste of time because there were few classes and to get involved in a process of discussion on the criteria implied a slow, drawn-out process, as is reflected in some research (Authors, 2013).

According to Boud and Falchikov (2006) universities should emphasize assessment for long-term learning focusing on authentic tasks in order that the students can face real problems related to everyday life and work. In this sense, an important factor to bear in mind is that the students take this unit before the Practicum and they have to apply the knowledge gained from the unit during the Practicum. The final assessment for this unit is done only after the students have had their Practicum experience when they have to apply the theoretical concepts they have learned to resolve real problems. As the tutors noted in the interviews: 'the practicum enabled the students to apply their classroom learning to real contexts and real problems, leading to very meaningful learning and what we can describe as authentic assessment'. This task could be accomplished with the key elements of authentic assessment (previously

named authentic achievement by Newmann and Archbald (1992) in Cumming and Maxwell, 1999), that is:

‘It should involve constructive learning, disciplined enquiry and higher-order thinking and problem solving. It should also have a value dimension, of aesthetic development, personal development or usefulness in the wider world. The last of these implies transfer of learning’ (Cumming and Maxwell, 1999:179).

William (2000), in his presentation for the European Association for Educational Assessment (as quoted in Kirk and O’Flaherty, 2003), wrote that if we want students to be able to apply their knowledge and skills in new situations, to be able to deal with relatively unstructured problems and to reflect on their own performance, then the tasks they undertake to assess their learning must require them to demonstrate all of these attributes. The current study indicated that the target unit fulfilled all of this. As shown above, during the Practicum students were required to apply their knowledge to real situations through a variety of tasks and activities.

It is also important to highlight how the school-based practice provided an excellent opportunity for self-assessment as the students were constantly aware of the standards required of them and this acted as a very valuable tool for analysing their own performance. As James, one of the tutors, said:

‘At the end of every class they taught they had to answer a series of questions which helped them to reflect on their performance and on whether the session objectives were achieved. In every class they had to note down issues to consider when they are preparing the next session and which aspects of their own teaching skills they wanted to improve’.

During the Practicum students were supervised by a mentor from within the Practicum school who gave them feedback on how they could reach the required standards of professional performance. At the beginning of the subsequent year and

before the start of the next Practicum, the students completed a new self-evaluation which identified their strengths and highlighted those areas they needed to continue improving during the following semester in order to achieve the required QTS standards. This self-evaluation was used solely as an exercise in self-reflection and as a point of reference for ongoing improvement, so it was not marked or graded. Published research supports the view that self-assessment, peer-assessment, co-assessment and negotiated assessment should be incorporated into the design and delivery of the units if we want to fully engage students in their own process of learning in order to promote the kind of dispositions they need to be lifelong learners (Carless et al., 2006; López-Pastor et al., 2006; Authors, 2014) and develop evaluative expertise (Sadler, 2010). Although these formats were permitted and encouraged by the unit specification they were not actually employed to any significant degree.

The fact that students applied the materials produced in their Teaching and Learning 1 unit (undertaken in the very first semester of the course) within the Assessing Learning and Developing Teaching unit can also be considered as AfL as this work was returned with additional constructive feedback from the tutor. This element was graded and had a section for formal feedback. As Ann, the other tutor of the unit commented: 'The students used the work they produced for the Teaching and Learning 1 unit as the basis for the assessments they prepared for the classes they delivered on the Practicum'. With regard to the first activity assessment-grading (1) it became apparent for the tutors that "it was very difficult to provide the students with feedback before they started their Practicum as there was no time for it"(Ann). However, students need to receive feedback at a time when they can use it for the purposes of current work and/or see how it feeds forward into future tasks (Carless et al., 2006). It may have been beneficial to resolve this issue so that the work could be corrected and the revised

version implemented during the Practicum. To achieve this, the work would need to have been submitted earlier or the feedback could have been submitted via the University's online virtual platform.

At the conclusion of the teaching of AfL in the unit, our reading of the data suggested to us that this unit acknowledged and was quite well aligned with the criteria for teaching and assessment as contained in the University Strategic Plan. It was connected with other units; balanced summative and formative strategies; the unit outcomes were consistent with the objectives as stated in its specification document; it provided feedback and self-regulation through self-assessment during the practicum; and AfL was largely used during the seminars. In general, the unit also reflected the type of assessment required according to the National Curriculum (2007). This latter was a vital factor, given the content of this unit, as the message needed to be both clear and coherent within the government's stance on assessment practice and was closely aligned to the concept of AfL. We could say that on the whole, the approach in the unit was learning-oriented assessment as is understood by Carless et al. (2006), in other words, tasks were 'fit for purpose', students were involved in the assessment process in ways which support the development of evaluative expertise, and feedback was forward-looking and could be acted upon (Carless et al., 2006, p.396).

Nevertheless, the unit leaders didn't ask for students' expectations at the beginning of the process, neither did they engage them in criteria setting; they didn't use peer-assessment and self-assessment for grading, just with formative intentions. As we observed over the process the tutors involved in the study retained a high degree of control over their students' work by showing them how they could improve their performance. However, from a formative perspective it could be interesting to see how students themselves would reflect on how to improve, with the support of their tutors.

Thus, shared or co-assessment could be used to achieve this (López-Pastor,2008; Authors, 2013). All of this indicated that students did not have a “voice”. In the next section, we consider whether this issue, among the others we have discussed concerning the teaching of the unit, influenced students’ conceptions of AfL.

Student-teachers’ understanding of AfL

Reflecting on our second research question, we wanted to know how effective were the unit leaders in communicating the concept of AfL to the students and what understanding of this concept the students developed. For this question it was necessary to analyse: the understanding of the students about this concept once they finished the unit and the importance and meaning they gave to this learning; to what extent they applied this learning in their Practicum, and whether they intended to incorporate this type of assessment in their professional practice?

Students’ constructions of AfL. All students interviewed were asked how they defined the concept of AfL and to state whether in their opinion this concept was applied to the unit. Our main findings are that for most of the students the concept was clear and matched with their teachers’ conceptions of AfL. Nevertheless none could provide a complete definition, which may be due to them having forgotten the exact concept but remembered the general idea from the unit the previous year. Furthermore, the majority of students had difficulties identifying differences between formative assessment and AfL, being more familiar with the latter. In addition, all students except three could identify the strategies used in the unit as AfL. Just two said no AfL strategies were used in the unit. And the whole group interviewed agreed that the content learned in this unit was useful and relevant to them as student teachers of physical education.

Beneath these headline findings, we will now discuss the results in more detail. The main feature of AfL the students teachers highlighted associated assessment with feedback delivered during the process in order to improve the pupils' learning. As John said, 'AfL is the assessment that is happening during a lesson or a unit. When teacher give FB to pupils'. Nevertheless, for Peter, Cathy and Rose the concept was not so clear. They showed some confusion about AfL in relation to criterion referencing and norm referencing types of assessment. For instance, Peter identified AfL as 'assessing individuals against themselves, not comparing with others'.

Regarding the difference between formative assessment and AfL, most of the students were more familiar with AfL. On the one hand, they were not able to identify differences between these two concepts. On the other hand, to know whether students understood the concept it was interesting to find out if they could identify the unit assessment strategies as AfL. In this sense, Peter, Cathy and Rose could not find examples or didn't have clear idea. They were the same students that appeared not to have clearly grasped a concept of AfL.

In comparison to these three exceptions, the rest of the students did identify the strategies used in the unit as AfL. But highlighting differences between lectures, seminars and placement, John noted, 'Maybe. There was feedback during the unit; in the seminars there was more interaction with the lecturer. I received feedback'. Even, Mary and Susan criticized the fact that they didn't receive FB about the assignment 'Yes, in the seminars, in the tasks we do. But the assignment wasn't formative because I didn't receive FB about the assignment before going to placement. No FB during the process but yes in the placement' (Mary); 'Not exactly, because assignment was summative, not formative. During the lessons was formative through feedback' (Susan). In this case, the students identified very well the concept and went further, showing that

what was most important and useful for them was the feedback rather than the grade. This continuing use of summative assessment may be evidence of some hesitation on the part of the tutors to fully commit to AfL in their own assessment practices.

Otherwise, Cathy and Rose stated that AfL was not used in the unit: ‘nowhere could I see how I was being assessed in this way’ (Rose) which could show, as one possibility, they didn’t catch the whole meaning of AfL. Cathy identified formative and summative assessment as the same thing: ‘because teachers only assessed at the end of the unit’ (Cathy). These students understand that if you only take into account the final assignment to grade a piece of work, without the possibility of improving it, this is not formative assessment, a view which could corroborate what Taras (2009) explained as a lack of alignment and coherence in the rationale of the theory, and contradictions involving concepts of formative and summative. But also it could indicate that students’ perceptions are different from their teachers’, as in MacLellan’s study (2001).

Concerning the importance and meaning of learning given by the students to the unit it can be noted that all the students said this unit was relevant and useful to them, especially because of its application in the school, ‘Yes, definitely. It has application in the school’ (John); ‘Very relevant for teaching in placement because it helps to assess pupils’ (Robert). The most useful things for them were to know the different ways of assessing learning especially from a formative point of view ‘The most useful thing is to know how to assess, to know different ways of assessing learning, not only at the end of the unit. Before this unit I wasn’t aware of the importance of assessing’ (Rebecca); the core task and the practical issues that allowed the application of different strategies in placement: ‘The most useful part was what we learned in seminars and how to perform on placement, Assessment charts’ (Sam).

All students noted the knowledge learned during the unit was relevant in relation to the Practicum, although the question was asked in general, referring to the whole unit and not the Practicum specifically. Nevertheless the students stated they had applied this knowledge in different ways during the Practicum, observing and giving feedback during the lessons, using different assessment tools like reciprocal cards, establishing criteria and questions and answers. From the examples they gave we can see they understood the concept: ‘Yes. Using reciprocal cards. Pupils performed and their fellow pupils marked against the criteria. My tutor was interested, not in grades but in helping students’ (Sam); ‘Yes, questioning/answering was useful to help pupils when they were struggling’ (Mary). Only John said he didn’t have a chance to assess in his school, and he just helped the teacher in his assessment, ‘No, I didn’t. There are few opportunities to assess in year 2 school experience. There are more in year 3.’

Students’ intentions to incorporate this knowledge in their future professional practice.

Perhaps one way to gain an insight into the possible nature of future assessment practices of practitioners is to analyse the practices of pre-service teachers (Karp and Woods, 2008), insights which in turn could allow the refocusing of teacher education. In this sense, we only have data on what students teachers said about what they did. In future, research could usefully observe these practices directly and in depth.

We can note that all the students interviewed thought of using this knowledge as a practitioner but just two (Rebecca and John) said they felt confident they had learned enough about AfL: ‘Yes. It is quite an important theme and I think I have enough knowledge to deliver it to pupils’ (Rebecca); ‘I don’t need more knowledge, it’s quite enough’ (John). Most students felt there was still a lot to learn. As Cathy and Rose said, more practice was needed in other areas: ‘I feel confident in the games areas but I need

more practice with core tasks and criteria in other subjects' (Cathy); 'Yes, but it would be interesting to discover more areas where to apply this kind of assessment' (Rose). Also, Eli noted that it would be good to learn more about peer assessment, 'Yes, I feel confident in applying these ideas but would like more information and experiences about peer assessment. The most difficult thing is to know what kind of feedback each individual needs'. And Susan showed some lack of confidence, because she was still learning and becoming more aware of the importance of assessment: 'To feel confident you need the mentor's support, it's a key point because assessment is the most difficult part of teaching. You need to compare your marking with the mentor'.

Conclusion

Our purpose in this paper was to address two questions, the first concerned with the extent to which the concept of AfL employed in a physical education teacher education course complied both with university policy and the literature, and the second, the extent to which students' notions of this concept were consistent with their lecturers' understanding, with their university's policy, and with the research literature. In the case of the first question, we found that the unit was reasonably well-aligned with the university's strategic plan for teaching and learning, and crucially with notions of AfL within the National Curriculum in England and Wales. This alignment between national and university policy and the pedagogy of the PETE unit provided strong legitimacy for AfL as an alternative practice in physical education. In terms of the second question, despite the fact that they lacked a voice in framing their understanding and uses of AfL in advance of the unit, our data suggest that practically all the students learned the concept, not just as a theoretical framework but applying different strategies during their practice in the school. Perhaps due to the strong sense of legitimacy of the concept,

they found this knowledge relevant and useful, they recognised the need to learn more about it, and they intended to incorporate this knowledge in their future professional practice.

Although student teachers in this research showed their intention to use AfL in their future professional practice, recent research in the field of assessment in physical education tells us something about the barriers these novice teachers might expect to face in realising their intentions since alternative assessment practices are not in widespread use (see Hay and Penney, 2013; Lopez-Pastor et al., 2013). While Lopez-Pastor et al. provided evidence of genuine progress, they nevertheless were not complacent about the place of assessment in school physical education (López-Pastor et al. 2013).

We might conclude cautiously then that AfL is strongly defended in policies and the literature, and that the arguments about its validity and educational relevance have been more or less won in this context. But this does not mean that this acceptance of approaches to assessment such as AfL will also automatically transfer to the sphere of school practice. So our question to guide future research is, what happens to well-supported concepts and good intentions once the student teachers start to teach? Alternative forms of assessment have, after all, been advocated in the literature for over 30 years or more, but assessment in school physical education has not changed dramatically during this time. Perhaps these practices in PETE courses are not as widespread as we thought or there is no consistency between what tutors say should be done and what teachers actually do? According to Karp and Woods (2008), if shifts in teachers' beliefs are to be made, critical and authentic assessment experiences in teacher preparation will need to be maintained and carefully planned for within PETE programs. Also, as Penney et al. (2009) argue, even better alignment between

instruction, curriculum and assessment will assist teacher educators and teachers as well as the recognition of ways in which contemporary contexts of higher education (and associated with that, policies relating to assessment, to student recruitment and retention, and to broader performance expectations of academic staff) may limit/constrain staff in their ability to develop or strengthen the extent to which their assessment practices align with AfL. Also, it can give value in research that therefore identifies barriers that PETE staff face in seeking to extend and enhance their assessment practices. Along with this continuing effort in the policy and teacher education spheres, we propose that more research is needed on all participants' perspectives, especially on what student teachers understand and how they apply these concepts in the practice. Finally, if AfL is to become an embedded, routine and productive contributor to teaching and learning in physical education, we think we need to know more about the structures that support and reproduce (as Brown and Evans 2004 put it, intergenerationally) the prevalent and dominant practices of teaching and learning in school physical education that work against the use of AfL and related approaches to assessment.

Disclosure Statement: We acknowledge any financial interest or benefit is arising from the direct application of this research

Notes

ⁱ From Key Stage 3 National Strategy. Pedagogy and Practice. Unit 12: Assessment for learning. (2004).

ⁱⁱ López Pastor et al. (2009) in his book *Evaluación Formativa y Compartida en la Educación Superior* (Formative and Shared Assessment in Higher Education) has carried out a large review of the European Studies that show evidence about viability, validity, effectiveness and pedagogical value of the students' participation in the assessment processes at university. Those studies suggest the importance of the use of formative and shared assessment models for its coherence with the EHEA and for the advantages found.

ⁱⁱⁱ Please see: <http://nationalstrategies.standards.dcsf.gov.uk/secondary/assessment>

^{iv} General Certificate of Secondary Education (GCSE) is an academic qualification awarded in a specified subject, generally taken in a number of subjects by pupils aged 14–16 in secondary education in England, Wales and Northern Ireland. The General Certificate of Education Advanced Level (short form: GCE Advanced Level), or more commonly, the A Level is a school leaving qualification offered by educational bodies in the United Kingdom and the British Crown dependencies to students completing secondary or pre-university education. The AS Level is a qualification in its own right, and the AS Level combined with the A2 Level forms the complete A Level qualification.

^v The themes that emerged from the students' interviews data were related to: Concept of AfL; examples of AfL in the unit studied and in others; relevance and meaningful of learning on the unit; transference to real contexts (in the practicum) and intention to use these strategies in their future professional practice. After that, clusters of interviewees whose responses were similar to each other were done. These themes helped to structure the results section with two main headings: Students-teachers understanding of AfL and students' intentions to incorporate this knowledge in their future professional practice.

^{vi} Assessment has to be diagnostic, inclusive, connected with other units and balance summative and formative strategies. Assessment tasks have to be aligned with learning outcomes, clarifying assessment expectations, engaging them in criteria setting; providing feedback and self-regulation through participative strategies of assessment as self-assessment and peer-assessment.

^{vii} Sorbonne Joint Declaration (1998) and Bolonia Declaration (1999) agreements for a common space for European Higher Education Area (EHEA).

^{viii} According to Blanchard: 'When tutors or leaders present things clearly the culture of the class or the school tends to be transparent. When the pupils or staff do things for themselves and with others then the culture tends to be interactive.' (2009: 7)

References

Assessment Reform Group (ARG) (1999) *Assessment for Learning: beyond the black box*. Cambridge: University of Cambridge, Faculty of Education.

Black P and Wiliam D (1998) *Inside the black box: raising standards through classroom assessment*. London: Kings College.

Black P, Harrison C Lee C Marshall B and Wiliam D (2002) *Working inside the black box: assessment for learning in the classroom*. London: Kings College.

Biggs JB (1999) *Teaching for quality learning at university: what students does*. Buckingham: Open University Press.

Blanchard J (2009) *Teaching, Learning and Assessment*. Mc Graw Hill, Open University Press.

Boud D (1990) Assessment and the promotion of academic values. *Studies in Higher Education* 15(1): 101–111.

Boud D (2010) *Assessment 2020: Seven Propositions for Assessment Reform in Higher Education* (Sydney: Australian Learning and Teaching Council). Available at: <http://www.olt.gov.au/resource-student-assessmentlearning-and-after-courses-uts-2010>.

Boud D and Falchikov N (2006) Aligning assessment with long-term learning, *Assessment & Evaluation in Higher Education* 31(4): 399–413.

Braun V and Clarke V (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology* 3 (2): 77-101.

Brown D and Evans J (2004) Reproducing Gender? Intergenerational Links and the Male PE Teacher as a Cultural Conduit in Teaching Physical Education. *Journal of Teaching in Physical Education* 23, 48-70.

Brown S and Glasner A (1999) *Assessment matters in Higher Education: choosing and using diverse approaches*. Buckingham: SORHE and Open University Press.

Carles D, Joughin G and Mok M (2006) Learning-oriented assessment: principles and practice. *Assessment & Evaluation in Higher Education* 31(4): 395–398

Carless D (2010) Prospects of implementing assessment for learning. *Assessment in Education: Principles, Policy and Practice* 12 (1): 39-54.

Cumming JJ and Maxwell GS (1999) Contextualising Authentic Assessment *Assessment in Education: Principles, Policies and Practices* 6(2): 177-194.

DCSF/QCA (2007) *The National Curriculum - Statutory Requirements for Key Stages 3 and 4*. London: HMSO.

Declaración de Bolonia sobre el Espacio Europeo de Educación Superior (1999). Available at: http://magno.ub.es/fas/piune/normativa/declaración_bolonia.pdf

DfEE/QCA (1999) *Physical Education. The National Curriculum for England*. London: HMSO.

Fallows S and Chandramohan B (2001) Multiple approaches to assessment: Reflections on use of tutor, peer and self-assessment. *Teaching in Higher Education* 6(2): 229–246.

Fullan M (1991) *The New Meaning of Educational Change*. London: Cassell.

Glaser BG and Strauss A (1967) *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Goetz JP and LeCompte MD (1988) *Etnografía y diseño cualitativo en investigación educativa* [Ethnography and Qualitative Design in Educational Research]. Madrid: Morata.

Hay P and Penney D (2009) Proposing conditions for assessment efficacy in physical education. *European Physical Education Review* 15(3): 389-405.

Hay P and Penney D (2013) *Assessment in physical education: A socio-cultural perspective*. London: Routledge.

Karp GG and Woods ML (2008) Preservice Teachers' Perceptions About Assessment and Its Implementation. *Journal of Teaching in Physical Education* 27(3): 327-346.

Kvale S (1996) *Interviews An Introduction to Qualitative Research Interviewing*. Sage Publications.

Kirk D (2010) *Physical Education Futures*. London: Routledge.

Kirk D and O'Flaherty M (2003) Learning Theory and Authentic Assessment in Physical Education. Paper presented to the Physical and Sport Education SIG at the Annual Conference of BERA. Edimburgh, September 13-15.

Lincoln YS and Guba EG (1985) *Naturalistic inquiry*. Beverly Hills: Sage

López-Pastor VM (2008) Implementing a Formative and Shared Assessment System in Higher Education Teaching. *European Journal of Teacher Education* 31(3): 289-307.

López-Pastor V (coord) et al. (2009) *Evaluación Formativa y Compartida en Educación Superior. Propuestas, técnicas, instrumentos y experiencias*. Narcea: Madrid.

López-Pastor VM, Kirk D, Lorente-Catalán E, MacPhail A and Macdonald D (2013) Alternative assessment in physical education: a review of International literature. *Sport, Education and Society* 18 (1): 57-76.

Lorente E and 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.

Lorente E and Kirk D (2014) Making the case for democratic assessment practices within a critical pedagogy of Physical Education Teacher Education. *European Physical Education Review*, 20 (1): 104–119.

MacPhail A and Halbert J (2010) 'We had to do intelligent thinking during recent PE': students' and teachers' experiences of assessment for learning in post-primary physical education. *Assessment in Education: Principles, Policy & Practice* 17 (1): 23- 39.

MacLellan E (2001) Assessment for Learning: the differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education* 26(4): 307-318.

Moynihan C, Murphy G and O'Flaherty M (2006) Future designs – planning for a quality physical education experience, paper presented at British Educational Research Association, University of Warwick, UK, September 6–9.

Penney D, Brooker R, Hay P and Gillespie L (2009) Curriculum, pedagogy and assessment: three message systems of schooling and dimensions of quality physical education. *Sport, Education and Society* 14 (4): 421-442

Sadler DR (2010) Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education* 35(5): 535–550.

Sanmartí N (2007) *10 ideas clave. Evaluar para aprender {10 key ideas. Assessment for learning*. Barcelona: Graó.

Scottish Government (2011) Curriculum for excellence, building the curriculum 5: A framework learning. Edinburgh: Author.

Sluijsmans D, Dochy F and Moerkerke G (1998) Creating a learning environment by using self-, peer- and coassessment. *Learning Environments Research* 1(3): 293–319.

Stake RE (1998) *Investigación con estudio de casos* [The art of case study research]. Madrid: Morata.

Taras M (2007) Assessment for learning: understanding theory to improve practice. *Journal of Further & Higher Education* 31(4): 363-371.

Taras M (2009) Summative assessment: the missing link for formative assessment. *Journal of Further & Higher Education* 33(1): 57-69.

Veal ML (1988) Pupil assessment issues: a teacher educator's perspective. *Quest* 40 (2): 151-161.

Zabalza M (2001) *La enseñanza universitaria. El escenario y sus protagonistas*[*University teaching. The scenary and their characters*]. Madrid: Narcea.