UNIVERSITY OF BIRMINGHAM

Assessment in Maltese Physical Education: An insight on assessment practices in early secondary years

By

Karl Cortis

Dissertation presented to the University of Birmingham

In Fulfilment for the Degree in

MSc (Res) Sport Exercise + Rehabilitation Sciences

Word Count (27663)

December 2018

Declaration of Authenticity

I confirm that this research is my own work, not copied from any other person's work (published or unpublished), and that it has not previously been submitted for assessment on any other course.

Karl Cortis

Abstract

Purpose: The purpose of this research project, therefore, was to pursue this line of inquiry by exploring the nature, range and efficacy of assessment practices in physical education in the Maltese context. More specifically, the following research questions were addressed: (i) what are PE teachers' understandings of the notion of assessment and assessment policies in Malta? , (ii) what assessment practices / approaches do PE teachers employ and why? , and (iii) what barriers do PE teachers encounter that hinder their assessment practices?

Methods: The study was conducted in two phases. The first phase focused on examining PE teachers' perceptions and practices about assessment. A total of eight PE teachers agreed to participate in the first phase the study. The primary data collection tool was semi-structured interviews supplemented with lesson observations for the purposes of triangulation. Results from phase one informed the development of a national teacher survey (phase two). This was designed to explore teachers' perceptions and understanding of the purpose and application of assessment on a larger scale. Important questions around teachers' experiences and views on their initial teacher education and professional development opportunities in relation to assessment were also posed. The survey was distributed to all secondary PE teachers employed in state and church schools nationwide, and returned by 90, which represents a 71% response rate.

Results: In this study, teachers seemed to share very similar ideas on the multi-dimensional nature of physical education – i.e. they identified four domains in terms of what the focus and purpose of PE should be. This seemed to transpose onto assessment as assessment was overall understood to be used to assess not only the physical dimension by most teachers. It was also evident that the extent of the teacher's experience (i.e. years in the profession) has a direct impact in the mode of assessment of certain domains. Some concerns on the effectiveness of the current assessment policies and practices were highlighted, including questions about the nature and quality of teacher preparation to be able to assess in effective and meaningful way. However, even though teachers reported following similar syllabi, appeared to share most overall common perceptions about the importance and mode of

assessment, including its potential to affect student motivation, , the way they conduct assessment seems to vary. Finally, teachers also expressed some concern regarding the current assessments' focus on performance rather than progress vis-à-vis national policies.

Conclusions: Results reported in this thesis suggest that there are a number of important areas which need to be addressed. Firstly, it is important that PE teachers are aware of and acknowledge the existence of a wide range of assessment processes and procedures, including summative and formative assessments. It is particularly important that they share a more in-depth understanding of the importance and complexity of AfL. To address this, emphasis should be placed on the content and quality of initial teacher education and CPD as it appears that they currently neglect this important aspect of teaching and learning.

Keywords: Assessment; Physical Education; Assessment for learning; Formative Assessment

I dedicate this study to my parents Laurence and Miriam my first teachers. They were the first people to introduce me to the world of education, showing me the importance of assessment, evaluation and reflection for the purpose of self-improvement and betterment of the society we live in. I am truly grateful to you for having given me life and supporting me in it.

Acknowledgements

I would like to express my sincere gratitude to my tutor, Dr. Kyriaki Makopoulou for her academic guidance, patience and time throughout this MSc programme. Without your support and guidance, I could not have completed this venture.

My heartfelt gratitude also goes to my girlfriend Jasmin for the patience and care, and for all the times you were there!

I am also grateful to my colleague Ms. Charlotte Schembri who so generously proofread this dissertation.

Contents

Chapter 1. Introduction and study purpose

Within education, 'assessment' is most commonly understood as a means of determining student achievement in relation to the learning outcomes of the course (Nitko and Brookhart, 2007; Rink, 2010). The contemporary understanding of assessment, however, points to a dual purpose: (i) to monitor student achievement and progress, often referred to as summative assessment; and (ii) to facilitate and inform student learning (Huba and Freed, 2000; Lund and Tannehill, 2010). The latter is also known as formative assessment or assessment for learning (AfL). AfL serves broader pedagogical purposes and "refers to all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (Black and William, 1998, p.8). Over the last two decades, the consistent message resulting from relevant studies is that AfL should be embedded in teachers' practices, taking place on an ongoing basis and across a unit of work (CERI, 2008; Black et al., 2003).

Whilst research evidence into assessment in schools is accumulating in international literature (Hardman, 2011; Van-Vuuren Cassar, 2011), research into teachers' assessment practices is scarce in the Maltese educational context. One of the few published studies (Grima and Chetcuti, 2003) sought to shed light on assessment practices in Maltese schools across all the subjects. The research showed that both summative and formative assessment practices were commonly used, with the half yearly and annual periods being the most common times of assessment. The study also suggested that record keeping was not consistently employed. Since then, there was very limited published research about teachers' perceptions and practices about assessment in Malta. Moreover, in the context of teaching

PE, there is no research directly examining Physical Education (PE) teachers' assessment practices. Very little is therefore known about the way assessment is conducted in schools and the extent to which it is effective.

In 2012, a new National Curriculum Framework (NCF) was introduced. This brought about some fundamental changes in the ways teachers were expected to employ assessment. Alongside recognition that more time should be allocated to PE (NCF, 2012, p 18), as well as a clear emphasis on student-centred pedagogies (including inquiry based and cross curriculum opportunities), there was a clear shift in emphasis which highlighted the importance of both the process of assessment as well as its importance for measuring academic achievement and progress. There was a clear expectation, in other words, that teachers should use a range of assessment tools on a continuous basis with the aim to promote pupil learning. This change brought about by the NCF offered a clear research opportunity, since there is a lack of robust data examining whether and how secondary school PE teachers in Malta are practising assessment in line with these new curriculum requirements.

The purpose of this research project, therefore, was to pursue this line of inquiry by exploring the nature, range and efficacy of assessment practices in physical education in the Maltese context. More specifically, the following research questions were addressed:

1. What are PE teachers' understandings of the notion of assessment and assessment policies in Malta?

2. What assessment practices / approaches do PE teachers employ and why?

3. What barriers do PE teachers encounter that hinder their assessment practices?

This study provided insights into the existing situation in Malta with the aim of informing policy and practice. Understanding teachers' practices will help guide policymakers and teacher trainers in making sure that the correct resources are provided, and that the creation and implementation of new assessment policies are realistic and effective. All this is to benefit both the learners as well as the teachers who have the role of implementing assessment policies.

Chapter 2. Literature Review

This chapter aims at giving an overview of existing literature on assessment, delving into different formats and types of existing assessment practices. It focuses on the importance and value of assessment in the educational system and its role as an integral part in the process of teaching and learning. This chapter gives a structured view of the uses and perceptions of assessment and investigates the Maltese educational context, highlighting a gap in research in the current Maltese literature on assessment, especially in Physical Education.

2.1 What is assessment?

Assessment is one of the central responsibilities of educators, a significant aspect of teaching and learning (Hussain, Said, and Hilal, 2014). Assessment is a valuable tool that can be used to evaluate a programme of physical education as it can be used in order to collect evidence on what the students would have achieved in the course of the programme, scheme or unit of work (Avery, 2012). It has been argued that, overall, teachers spend a lot of time on assessment (Butt, 2010), however, the concept of assessment is broad and entails different dimensions. Thus, it is important to define what assessment is. Assessment can be said to be the process of collecting facts and data to understand students' progress and to use this information when planning for future 'educational course of action' (Mohan R. , 2016, p. 47). Without assessment, no feedback can be provided to neither teacher nor student (Trauth Nare and Buck, 2011). The optimal assessment provides a 'window' into the way a student thinks and understands a concept, as well as into the 'strategies' the student uses as a solution to a task (Schleicher, 2018, p. 277).

As previously noted, there are different types of assessment with a distinctive difference between what is known as formative (or assessment *for* learning) and summative (assessment *of* learning) assessment. The former is an ongoing process (Redelius and Hay, 2012), aiming to obtain insight into the actual process of learning, as well as to understand where students are in terms of current knowledge and understanding. It is intended to provide valuable information to teachers to further support student learning. The use/application of assessment for learning, and the insight this provides, encourages teachers to adapt teaching in order to respond to students' learning (William , 2006).

Summative assessment, on the other hand, focuses on measuring whether the learning outcomes have been achieved (Stobart, 2008), confirming what students know, and are capable of doing a task in relation to curriculum outcomes (Gibbons and Kankkonen, 2011). Summative assessment may be described as assessment of learning, serving the purpose of evaluating students' performance (Earl, 2003). This is a process that has the primary goal of quantifying what students know in 'a specific area of learning at a particular time' (Harlen W. , 2004, p. 1). For any assessment endeavour to work, it is important that clear, objective, and transparent criteria – or 'expectations and standards' are set (Nicol and Macfarlane, 2006, p. 201); and that this set of criteria and expectations of performance are shared with colleagues, pupils and other key stakeholders in the education process (e.g., parents). Once criteria are established, teachers need to collect data in a systematic and reliable way to show 'how well the learner matched the standards and expectations' (Angelo, 1995, p. 7).

Moss (2013) states that there is a dependency on the quality of assessment and capabilities of the assessor for accuracies on the summative assessments. Summative assessment can

directly influence the students' achievement, attitude, motivation and effort (Rodriguez, 2004).

While formative assessment is used at a lesson level, summative assessment serves the aim of showing what level the students have achieved in the unit, or goals being assessed (Mercier and Dolittle, 2013). Although not granting an opportunity for improvement might be a negative consequence of summative assessment (Earl, 2003), Lund and Kirk (2002) emphasised the importance of also using summative assessment as a tool to communicate with parents and other stakeholders the importance of participation in physical activities, through a log handed to parents.

In physical education summative assessment will usually take place at the end of the topic or term (Constantinou, 2017). Constantinou (2017) compiled international literature on the topic and concluded that in this form of assessment, teachers either create summative physical education tasks or use pre-set tasks set by the education division, based on a number of criteria which allow the students to demonstrate specific skills learnt. These results are usually used to give a final grade to the students (ibid.). In a study on the assessment practices of different physical education teachers in 13 secondary schools, the majority of assessment practices were summative (Veal, 1988), however, recent times are seeing a shift towards alternative modes of assessment and formative assessment (Lopez- Pastor, et al., 2013).

2.2 Formative Assessment

There has been an increased interest in having strategies for assessment for learning (Ministry of Education and Employment, 2012), (Employment M. o., 2014), (Employment M. f., 2015). Many researchers state that a positive effect can be left on learning when formative

assessment is used, although this is seemingly based on limited evidence. It is also unclear whether or not the variety of practices that are employed during formative assessment do indeed serve to hinder or enhance its implementation (Bennett, 2011). Formative assessment would have no or limited effect on its own without correct implementation (Hendriks, Scheerens and Sleegers, 2014). For AfL to be successful for student learning, it must be implemented well (Heitink, et al., 2016)

2.2.1Assessment for Learning (AfL) strategies

In 1989 an Assessment reform group was set up in the UK promoting assessment for learning (McDowell, Sambell and Davison, 2009). Klenowski (2009), suggests that a main approach to formative assessment, which is of frequent ongoing practice in classrooms, is Afl. This is viewed as an event that keeps the quality of the process of learning in focus socially and contextually (Stobart, 2008). William (2011) makes reference to the United Kingdom Assessment reform group which advocated the five components that should be in place for assessment to improve learning. These are: (i) provide effective feedback; (ii) involve students in their own learning, by having a clear understanding of where they are and what they need to do to improve further and making sure the learning intentions, goals and success criteria are clear and shared with the learner(William and Thompson, 2007); (iii) make effective use of assessment outcomes / insights to inform and adjust teaching practices; (iv) realise the impact that assessment has on the motivation and self-esteem of a student, therefore playing an important role in learning; and (v) involve students in self- and peer assessment (Wiliam, p. 39) therefore helping students become responsible for their own

learning through self-assessment (William and Thompson, 2007). Later in 2007 the assessment for learning group identified other key strategies for implementing Afl:

1. Gathering evidence of student learning through assessments, formal or informal (e.g. questioning or discussions);
2. Giving formative feedback to the learners as a means to support learning;
3. Providing the students in becoming partners in instruction by means of peer assessment and discussions (ibid).

As evident in these components, AfL placed students central in the assessment process (Elwood and Klenowski, 2002), by having opportunities to assess themselves and their peers. There is also greater emphasis on student autonomy, a vital aspect to support their learning (Black, et al., 2006).

All strategies are focused on the theme of using assessment to inform learning, although their interpretation has been varied with researchers emphasising that to reach the aims of AfL it is important that a there is deeper understanding and application of the principles mentioned (Pedder and James, 2012).

Hargreaves (2005) argues that there are two different approaches to use AfL, namely, measurement and inquiry perspective. From a measurement perspective, quantitative evidence is gathered about a student in a formal way on whether pre-set levels of achievement have been reached. On the other hand, from the inquiry perspective qualitative, information through observations and conversations is acquired to give feedback as a process of reflective practice for discovery and understanding (ibid.). Both Hargreaves (2005) and

Wyayy-Smith, Klenowski and Colbert (2014) agree that quality implementation of AfL needs the adoption of the inquiry approach.

An integral part of Afl strategies is feedback as this is clearly part and parcel of the learning process, where the link between assessment and learning is created and supported (Elwood and Klenowski, 2002). Hattie and Timperley (2007) state that feedback can be said to be a 'consequence of performance', and that is, information gathered regarding the aspect of 'one's performance or understanding' (p. 81) with the agent provider being the teacher or parent. Feedback can take various forms and can be done through verbal feedback by an agent, self-assessment or peer-assessment as discussed in the following sections

2.2.1.1Feedback

Feedback is not a new concept in education and physical education. For decades, PE teachers were made aware of its importance in supporting student learning (Lee, Keh and Magill, 1993). In recent years, feedback has been included as a fundamental AfL strategy (William and Thompson, 2007). For it to be effective, feedback cannot happen in a 'vacuum' , but it has to be given in the context where the learning is taking place (Hattie and Timperley, 2007, p. 82). These authors present a model of feedback the purpose of which is to reduce discrepancies between current understandings and the desired goal. This can be reduced both by the teacher and or by the students themselves. Hattie and Timperley (2007) continue to highlight increased effort by the students or lowering of goals, while suggesting that

teachers have to provide specific and appropriate goals, suggesting that feedback has to answer three main questions: 'Where am I going?',' How am I going?',' Where to next?'; with each question working on four levels: task, process, self-regulation and self-level (p.87). Although feedback can be said to be influential on learning and achievement, its impact can be 'either positive or negative' (ibid., p. 81).

Good feedback needs to deliver specific information in relation to the tasks and processes of learning, filling the gap between what has to be learnt and what is actually learnt. (Sadler, 1989). Feedback is so crucial that a 'test' given a formative manner is only successful in its formative purpose if the teachers use its results to actually 'give feedback', and not only to give a grade to the students (Dunn and Mulvenon, 2009, p.2).

Hattie and Timperley (2007) continue to say that feedback is not only given by the teacher or instructor but it can also be given through parents, peers, books and also from their own selves (ibid.). All in all, the main aim of feedback is to encourage students to narrow the gap in their learning. The Australian Institute of Teaching and School Leadership (AITSL) highlights the importance of the students being able to realise the learning intentions and outcomes. Feedback aims to help the learner set his or her own goals while tracking their own performance when compared to the set goals. It can be said to be mainly used for evaluation and direction (Hattie and Timperley, 2007).

2.2.1.2 Self-Assessment

Self-assessment can be defined as the students evaluating their personal learning progress (Andrade and Valtcheva, 2009). It is an essential tool in assessment as through this type of self-assessment the students can identify their weaknesses, know where they need to focus

to improve, set their own goals and keep track of their own progress (ibid.). Self-assessment requires the learner to be able to recognise when the work done is correct, and also to arrange and adapt performance when it requires correction. It revolves around the notion of taking action to improve performance, as well as reflecting on what is currently being done. It is a tool that helps develop the learners' intuition to improve as a learner (Claxton, 1995).

One important precondition for self-assessment to work is for students to be motivated (Ćukušić, Željko and Jadrić, 2014) and educated (McConnel, 2006) in how to engage with the process. In other words, teachers need to invest time to support students to understand on how to assess their own learning (Andrade and Valtcheva, 2009). There is some evidence to suggest that it is possible for students to actually be capable of assessing themselves accurately when the assessment is not part of the summative evaluation (Darrow, Johnson and Williamson, 2002). In self-assessment the learner is helped by the teacher to reflect, evaluate and think upon the performance, thus realising the strengths and weaknesses. In this kind of formative discussion with the teacher in the role of a mentor, the learners can then plan correctly according to the knowledge achieved, thus becoming masters of their own learning (Claxton, 1995). Initially this could create a bit of confusion as to whether the results given are reliable, but the aim is to help the learner in critically reflecting upon actions. This helps shift the authority of marking and evaluating to the learner (ibid.).

It has also been argued that effective application of self-assessment can lead to improved teaching and instruction as students reflect on their progress and discuss their goals with the teachers. According to Andrade and Valtcheva (2009), reference to criteria is a vital part of self-assessment. The AITSL also mentions the responsibility the teacher has to create a 'classroom culture', where the culture of feedback gives students the necessary tools to

take informed decisions on how to proceed in their learning. It is therefore important that every student understands what the criteria for success are (AITSL, p.13).

In physical education self-assessment does not only involve the students self-marking their fitness results (Monsterrat, 2007), but gives power to the students to make decisions in their own learning (Lopez-Pastor, et al. 2013). In their guidance regarding subject assessment, the Northern Ireland Curriculum Counsel (Assessment, 2007) suggests using video analysis for self-assessment, as well as other forms of assessment like bulls eyes sheets (Hatfield and Phillips, 1989) where the students are given sheets to mark their performance and personal qualities, as well as Open Boxes (Carroll, 2005) where the students comment on their performance and areas of improvement.

2.2.1.3 Peer Assessment

A definition of peer-assessment would be that the students learn to give feedback to each other, while also developing tolerance to criticism (Claxton, 1995). Sometimes peer assessment solves the problem of the students receiving 'too little or no feedback from the teacher' due to the size of the class and the limited time given (Gibbons and Kankkonen, 2011, p. 8). Peer assessment also provides a great opportunity to receive immediate feedback (Johnson, 2004). In peer assessment the assessor is taking up the role of the observer and instructor. The student in this role can give individualised feedback, helping him to understand the skill on a cognitive level. One may therefore say that the benefit of peer assessment is positive to both observer and the student doing the skill. Gregory, Cameron and Davies (1997) highlight the importance of having clear criteria, and a very close resembling representation of what a good performance of the skill actually looks like. Gibbons

and Kankkonen (2011) state that in peer-assessment the role of the teacher is twofold: creating and providing clear criteria of assessment, and helping the observer by providing enough opportunity to observe, and be guided in observing the right things.

For reliability and validity's sake, the teacher must ensure that the skills chosen are reasonably easy for the beginner assessor to observe. Gibbons and Kankkonen (2011) further add that if there is too much complexity in the skill to be observed, then the student assessor will not be able to give good and valid feedback. This issue in peer assessment was also noted by Andrade and Valtcheva (2009), who argued that it is vital to coach students in how to assess, and also by providing them with clear criteria on what should be assessed. These authors insisted that the students should be given enough practice in assessing prior to official peer assessment.

2.2 Use and perceptions of Assessment

The evidence collected via formative and summative processes should consequently be used for various purposes. These include: (i) to provide students with feedback on their progress, a valuable process that can help both teachers and students to make timely interventions to improve (Mohan R. , 2016); and (ii) to provide the educator with an understanding of the effectiveness of their practices and the relevance of the curriculum (ibid.). This can subsequently inform teachers on what needs to change to support student learning further (Harlen and Crick, 2003). The strength of formative assessment is that information is gathered all the time so 'timely' pedagogical changes can occur and further support offered to aid pupil learning. Both purposes form an important part of student development and learning (Black and William, 1998). William (2006) considered the ability of

the teachers to be able to adapt their instruction to meet the needs of their students, as a very high achievement for the educator.

Long before the development of the concept of assessment for learning, there were publications acknowledging that assessment can take place at different points in time during any educational endeavour. Researchers had identified four distinctive points when assessment becomes even more pertinent. According to Airasian and Madaus (1983, cited in Mohan, 2016, p. 67) these are: Placement evaluation, where an evaluation of the entry level of a student is targeted (or what would be called today 'baseline' evaluation); diagnostic evaluation used to diagnose any difficulties the student might have hindering improvement; summative evaluation, where the degree of achievement is evaluated; and formative evaluation that is an ongoing process during the period of instruction.

As this outline of the four ways of assessment suggests, the amalgam of a range of assessment is what educators should strive to achieve rather than implement only one type of assessment. Mohan (2016) states that evaluation can play a significant role in improving student learning by making learning outcomes clear, offering short term targets to work for, giving feedback, and by obtaining information for the choice and selection of future learning experiences.

The literature on assessment in physical education echoes many of the issues, definitions and purposes discussed so far (Avery, 2012; Collier, 2011; Penney and Hay, 2008).Penney has recently argued that there is a lot of discussion on what should be the areas of assessment with regard to knowledge, skills and understanding in physical education (Hay and Penney, 2013). Although in some areas PE appears to be prominent in innovative practices (Penney, et al., 2009), it requires attention in each dimension including assessment and the 'linkage'

between assessment, curriculum and pedagogy (p. 430). Assessment is a social activity, required, developed and implemented by people (Hay and Penney, 2013). Assessment is the driving force behind changing pedagogies and curricula, with Lopez-Pastor, et.al. (2013) describing it as an ever arising issue which troubles physical educators, with the risk of students becoming 'grade hunters instead of knowledge hunters' (Togfors and Öhman, 2016, p. 164). To this end the Canadian Physical Education Curriculum framework highlights the importance of a multi-dimensional approach where assessment practices are done formatively, encouraging student summative evaluation and curricular development to focus on social responsibility, self-actualisation and the learning process itself rather than the grade (Department of Education, 1996).

Assessment occurs through three different media: pedagogy, curriculum and finally evaluation (Berstein, 1971). Hay and Penney (2013) place an emphasis on the importance of the PE teacher being literate in assessment, both in its implementation and also in the interpretation of its outcomes. Comprehension, application, interpretation and critical engagement are the four main elements of assessment literacy. Getting to know the barriers that teachers must overcome in assessment, and what practices are successful and why they will help in signalling some directions for the future of assessment in physical education (Lopez-Pastor, et al., 2013).

In line with the main points discussed in the previous paragraphs, physical education assessment can also be used to track the progress of the students, and to provide teachers with valuable information to adapt and adjust the method of instruction in order to meet the diverse needs of their students (Avery, 2012). Collier (2011) urges teachers to use assessment not as something separate from important decisions about teaching and learning but, in line

with Hay and Penney's (2013) theorising, as an integral part of the process of teaching and learning.

Overall, researchers in education and physical education acknowledge assessment as an important aspect of pedagogy and curriculum development (Black and William, 1998; Stiggins, et al., 2006). There is research evidence to suggest that teachers' perceptions might influence the mode of assessment employed (Ndalichako, 2015). Ndalichako (2015) has recently argued that when teachers perceived assessment as being there for giving feedback to improve their students' performance, they were more likely to use assessment as a guide to their teaching and in the learning process of the students. This, however, was not the case when teachers considered assessment as a tool to generate grades and reports (ibid., 2015). In this context, it is important to review some of the research available on understanding teachers' perceptions and use of assessment.

A study on the perceptions of secondary school teachers on assessment by Ndalichako (2015) , showed that general perception on assessment was quite favourable, although a number of teachers voiced the concern that assessment procedures would continue to add to their workload. In the said study, around 21% of the participants emphasised that the real purpose of assessment was to get the students ready for a final exam. Ndalichako (2015) argues that exams have a significant impact on the learning and teaching process, giving rise to the situation in which teachers make use of assessment as a means of preparing students for exams rather than to support the learning process. This conflicts with the idea that teachers do value formative assessment and give it importance in the promotion of learning by using information from formative assessment to support their teaching (Sach, 2012).

2.3 Teachers' perceptions and practices on assessment: Research in Malta

The available literature on the Maltese physical education scenario with a specific focus on assessment practices is scarce. It was only recently that a study was conducted (Falzon, (2016), shedding some light on this matter. This was a two-phase study, where a survey distributed to all secondary school Physical Educators was followed by semi-structured interviews. A stratified sample of twenty participants were chosen for the interview. Findings from this study suggest that there is a correlation between the mode of teaching and the mode of assessment employed. Specifically, teachers who opted for more conventional teaching approaches, such as command style of teaching, were more likely to rely on traditional methods of assessment. These included evaluating students against a set marking scheme through observation during the lesson, and giving a mark on what they remember about the student; a clear focus on assessment of learning – judgements of student performance. These also appeared to have a restricted understanding of assessment as they did not even consider assessment as an integral part of student learning. On the other hand, teachers who claimed to be using innovative pedagogical practices, such as peer teaching and guided discovery approaches, also claimed to observe an improvement 'not only the quality of their lessons but also increased students' participation, motivation, knowledge and attitudes towards physical activity'. These also had a more open approach to different methods of assessment, like peer assessment (Falzon, 2016, p. 70). The researcher of the aforementioned study felt that a deeper understanding of assessment types and what was really happening in Maltese physical education lessons was needed (Falzon, 2016).

2.4 The Maltese Context

Education in Malta is compulsory for all Maltese children from the ages of five till sixteen. Education is available in State, Church and private schools. Primary Education starts at the age of five and lasts for a total of six years. Secondary education begins when the students are eleven years of age. The latter is a five year course which leads to the Secondary Education Certificate (SEC). The study focused on this area in education, as it is the last course where physical education is compulsory as part of the curricular programme. Those who obtain the required number of SEC passes can opt to go to one of the sixth forms on the island which prepare them for the Matriculation Certificate which is required for entry into the University of Malta. A student may also opt to enter a technical institute, for which entry requirements vary depending on the course and the level.

2.4.1 Educational reform in Malta in the 20th and 21st century

Education in Malta is free for all citizens. Elementary education was free of charge from 1847, and declared compulsory in 1946. In the first Education Act in the 1980s it was explicated that all children and young people have the right to free education, underlying the state's duty to promote education and to ensure accessibility to all citizens. There is also clear reference to the important role of parents who have the duty of registering their children in schools and supporting them to attend regularly.

The 1988 Education Act also gave the state increased responsibility over what students should learn at school with the introduction of a National Curriculum Framework, which all schools should adhere to and implement. Physical education was a named activity in the Act making it one of the subjects to be taught in schools (Education Act, 1988). In the

subsequent revision of the national curriculum (National Minumum Curriculum, 2000), PE was identified as one of the main subjects in both primary and secondary education. This document acknowledged that teachers have an important role to play in supporting students to develop their physical, emotional, intellectual and social qualities and skills.

Before the year 2000, little emphasis was made on formative assessment. But this dramatically changed when the curriculum was revised in 2000 and which highlighted the importance of developing an array of abilities through the process of formative assessment (National Minumum Curriculum, 2000). The National Minimum Curriculum (NMC) was set out to strengthen the idea that assessment must be beneficial to students by suggesting that the information collected is carried forward from year to year to ensure that there is continuity on individual progress and individual development. In this way no individual is compared to norms that are universal – and can be excluding - in nature.

The 2000 national minimum curriculum gave great importance and emphasis on skills, competences, attitudes and values and not merely on the acquisition of knowledge and information (National Minumum Curriculum, 2000). It was argued at the time that its introduction was intended to bring a much needed change in the culture around assessment (ibid.). In fact one of the pillars of the National Minimum Curriculum (2000) was 'Principle 9: A more formative assessment' (p.29) where the focus would be on the individual student, helping in developing abilities by providing a more complete picture of the skills achieved. Looking back, the curriculum documents failed to specify in what ways this can be done, as no guidance was offered on how to do this in practice.

In 2012, another major revision of the national curriculum framework (Ministry of Education and Employment, 2012) was introduced, with a clear advocacy on the benefits and

importance of lifelong learning. A vision on supporting children and young people to love learning was also articulated: 'Success in promoting a love for learning and holistic development comes with suitable pedagogies to incorporate play and experiential, joyful learning' (Employment M. f., A National Curriculum Framework for All, 2012, p. 34).

Embedded in this revised curriculum is the need for schools to provide high quality curriculum provision to achieve 'a higher quality in the learning programmes and in the pedagogy with the scope of attracting learners to lifelong learning' (Employment M. f., A National Curriculum Framework for All, 2012, p. 31) . There were clear advancements as far as assessment is concerned too. For example, the role of assessment was described as being a tool to provide feedback to students, parents and teachers in a 'timely, and qualitative' way (ibid., p.41). A clear reference to the notion of 'Assessment for learning' (AfL) was evidenced and its importance underlined for the first time. AfL was described as a process and a method that takes place alongside 'learning' and this was a significant departure from the conventional views of assessment as something that takes place as a conclusive evaluation of what was learnt (ibid., p.42). In line with contemporary conceptualisations on the notion of assessment, it is clearly stated that the main focus of assessment would be threefold. Firstly, assessment should be used as a means to provide feedback about the teaching and learning pedagogies and methodologies employed. Secondly, assessment should serve the purpose of providing information about the learners' knowledge and abilities, rather than be used as a tool to stream them into classes of different levels as was customary up until 2011 (National Curriculum Framework, 2012). Thirdly, the notion of self-assessment was introduced for the first time and this was understood as an important aspect of assessment and learning, as students were expected to have increased involvement in their own learning and decision making (National Curriculum Framework, 2012). This emphasised the role of the teachers

where one of the main responsibilities of the teacher is to assess, record and report on the progress and development of the students assigned to him/her (Job Description Teacher).

It is also important to note that in 2018, a set of Learning Outcome Frameworks (LOFs) for each subject were issued to support NCF implementation. These learning outcomes set clear benchmarks for every subject on which assessment to and for learning was to be constructed upon. These were based on the eight key competencies listed in the framework by the European Union:

- Communication in the mother tongue
- Communication in foreign languages
- Mathematical competence and basic competences in science and technology
- Digital competence
- Learning to learn
- Social and civic competences
- Cultural awareness and expression
- Sense of initiative and entrepreneurship

These would be a foundation to the specific LOFs for physical education.

2.4.2 NCF and Physical Education

In the current 2012 National Curriculum Framework, the percentage of Physical Education lessons in the school time increased to 10% of the total curriculum time and the Physical Education teachers were presented with a new curriculum. The would mean approximately two hours of physical education lessons per week for year 7 and 8 classes. This had a set of learning goals in line with the NCF which stated that:

The learning experiences that take place during physical education activities aim at equipping learners with the necessary knowledge, competencies, skill, attitudes, and values which they need to maintain, promote and enhance physical, emotional, psychological and social wellbeing throughout their school life and as lifelong learners. (Ministry of Education and Employment, 2012, p. 35)

The learning goals for Physical Education represent the aim of physical education in schools and incorporates not only improvements in movement skills, but also engagement in decision making and problem solving. Students should also demonstrate a lifelong commitment to an active lifestyle and engage in ongoing evaluation for improvement (Handbook for the Teaching of Physical Education, 2012).

The content covered over the 5 scholastic secondary years is subdivided into 3 levels: Level 1 for years 7-8, Level 2 for years 9-10 and Level 3 for year 11. The purpose is to provide guidance on what is expected to be achieved during each year.

A student's final level is determined by the end of Form 5. Every year, the teacher allots 100 marks. The maximum number of marks attainable are 500 (100 x 5 years) which are translated into level 3. The following table explains which attainment level the student has achieved:

Table 1 Student attainment levels

Level 1	Level 2	Level 3
40% - 55% of 500	56% - 75% of 500	76%+ of 500

The major sections are focused on offering the students a holistic development where focus on transferable skills, observing and describing own performance, knowledge of rules and the ability to self-assess and self-measure physical and skill components. To date, only the Year 3 and Year 7 LOF's have been introduced in schools. For the purposes of illustration, the year 7 LOFs are presented briefly in the following paragraph.

The Physical Education curriculum mentions explicitly that the aim of Physical Education should be to develop the four main domains or strands, being the physical, mental, social and emotional domains, to be targeted by the areas of teaching mentioned subsequently It should also foster a positive attitude towards physical activity and sport in general, apart from developing skills that will allow the students to engage in a variety of sports, developing enough understanding of the activities presented to solve challenging situations and cultivating attitudes to maintain health and fitness throughout their lives.

To achieve the four aforementioned learning strands, in Years 7 (and 8), students are expected to have experiences in four main teaching areas. These teaching areas were structured on a research by Kenneth Hardman (2008) who explained that there was a shift in the PE curriculum, showing that its purpose was gradually being redefined as to include a wider spectrum of educational outcome which would include health related concepts. Hardman also noted that at the time of the study an emphasis was being done on team games and athletics, which was deviating from the current trends, minimising its relevance to the students effected (Hardman, The situation of physical education in schools: a european perspective, 2008)**.**

Teaching area one consists of *fitness*, and teachers are encouraged to provide learning opportunities that support students develop and improve their endurance, speed,

agility and coordination. Students should also develop their knowledge on basic training methods and exercises. The second teaching area consists of *team games*, particularly invasion and net games.

The third teaching area consists of *individual activities*. Named activities include athletics, swimming, gymnastics and educational dance, with outdoor activities making up the last area of teaching (Physical Education: Teaching Objectivesand Learning Outcomes, 2012). A degree of flexibility is evidenced as for each academic year, each college or school is responsible to teach four areas: Fitness, Team games and an individual Activity are obligatory, while the school can then opt to do outdoor activities, team games or an individual activity. For example a typical year 7 scheme of work would have Fitness, Handball, Sprints and Basketball with the year 8 scheme having fitness, football, Long jump and softball.

2.5 Technology in Assessment

Providing an authentic and accurate assessment experience that is up to the task in responding to the contemporary trends in physical education curriculum documents (Thorburn, 2007) presents a challenge on an international level. Embracing new approaches that include the learning types and opportunities that today's technologies bring with them is crucial (Penney and Hay, 2008). Literature has given attention to finding out how the current generation of students respond to digital technologies and any pedagogical concerns in relation to these technologies (Bennett, Maton and Kervin, 2008). There are a lot of ways to incorporate technology in Physical Education, - from sending emails to students and parents to enhancing the learning experience (Brooksie, 2014). A renowned website 'Support for real teachers' gives an list of technological tools one may use in physical education to keep

abreast with current trends (Apps for Physical Education, 2018). These are: Pedometers, Heart Rate monitors, Health tracking, Apps, video resources and games. Keeping abreast with current and developing technologies can be a challenging task for teachers. Sometimes Physical Education teachers can feel that the use of technology might not be applicable to their subject. Embracing technology can allow PE teachers to increase in the variety and dynamics of their classroom. Apart from this, technological aids can also be more appealing to the different students that make up the class (SPARK, 2016).

Chapter 3. Methodology

The purpose of this chapter is to outline the methods used to investigate the research that is the focus of this study. It also aims to give a rationale for why certain procedures were chosen, including the study design, the sampling process and size, the modes of data collection, the method of data analysis, as well as all ethical considerations. This will allow the reader to assess the study's credibility and transferability; as well as to allow other researchers to replicate the research project (Von Diether, 2016).

Drawing upon mixed methods, the present study was conducted in two phases. Phase one aimed at gaining in-depth insights into PE teachers' perceptions and practices about assessment. This involved an interview with the selected PE teachers and a lesson observation with a focus on capturing any assessment-related teacher behaviour for triangulation purposes. Results from phase one informed the development of a teacher survey which sought to explore teachers' perceptions and understanding of the purpose and application of assessment on a larger scale, their prior initial teacher education and professional development opportunities with a focus on assessment, and whether teachers see current assessment practices as being conducive to learning. The survey was distributed to all secondary PE teachers employed in state and church schools nationally.

Since there is evidence to suggest that educational surveys response rates are poor, and that schools tend to be over-surveyed (Sturgis, Smith and Hughes, 2006) , having interviews and observations to support the survey would be beneficial. Using this type of methodological triangulation would also provide a richer picture (Denscombe, 2007) of the data gathered.

Onwuegbuzi et al. (2011) describe this legitimation as having the 'weaknesses of one approach being addressed by the strengths from the other approach' (p.1261). Arguably, both qualitative and quantitative methods have their own strengths and weaknesses (Robson, 2011), so ultimately the mixed methods approach could merge and strengthen the research.

Since the researcher is a teacher by profession and having taught for nine years in the secondary sector, he experienced first-hand teaching PE classes in the secondary sector. He had hands on experience in assessment and its role in physical education and the effects it has on the process of teaching and learning. The researcher felt that it was imperative that this area in physical education is researched in depth. Therefore, it can be said that the epistemological stance would be more of an interpretivist one, interpreting certain elements of the study due to natural interest. The researcher tried to interpret the existing systems as described by the participants (Goldkuhl, 2012).

3.1 Research Design and sampling

In phase one, a case study design was employed. This decision to conduct teacher case studies was rooted in the purpose of the study; and that is, to develop an in-depth understanding of PE teachers' perceptions and practices about assessment. It has been argued that case studies, as a methodological tool, offer the potential of gaining access to deeper insights about contexts, situations and experiences (Eisenhardt, 1989; Sturman, 1999). Yin (1994) defined the case study as "an empirical inquiry that investigates a

contemporary phenomenon, the 'case', within its real-life context" (p. 13) in considerable depth. As case studies can seek to achieve different aims, it is important to clarify that the present study falls into the category of 'descriptive' case studies, with the primary aim of presenting an 'opportunity to observe and analyse a phenomenon' (Yin, 1994, p. 40), in this case being individual PE teachers working in secondary schools.

A total of eight PE teachers were selected as case studies, using purposive sampling with snowballing when required. Teachers recruited were chosen depending on the following criteria: Type of school in which they work (i.e. State or Church schools, co-ed or single education schools), and Gender and Career stage (newly qualified teachers and experienced, as well as male and female PE teachers at different stages in their career in order to gain cross experience and cross gender data). The final sample of the eight PE teachers who agreed to be study participants consisted of two male and two female teachers who were chosen from church schools, two being experienced teachers (>10 years in teaching), and two being NQTs or teachers with less than 5 years of experience. In the same way, another four teachers were chosen from State schools. Involving PE teachers working in both church and state schools was important to reflect the local school demographics (a total of ten state colleges and 26 church schools). Independent schools were not included in the study as they follow a different curricular setup in Physical Education.

The interviews were conducted between January 2018 and March 2018, and lasted between 34 and 50 minutes. They were conducted with each of the eight case study teachers at times and places convenient to them. Key features of the teachers' profiles can be summarised as follows:

Table 2: Criteria for interviews and observations

Type of School	Students taught	Gender	Experience	Code
State School	Co-Ed	2 Males	1 Male with more than 10 years of experience 1 Male with less than 5 years of experience	T1,M T2,M
State School	Co-Ed	2 Females	1 Female with more than 10 years of experience 1 Female with less than 5 years of experience	T3,F T4,F
Church School	Boys	2 Males	1 Male with more than 10 years of experience 1 Male with less than 5 years of experience	T5,M T6,M
Church School	Girls	2 Females	1 Female with more than 10 years of experience 1 Female with less than 5 years of experience	T7,F T8,F

In phase two, and building upon results from phase one, a cross-sectional design was adopted with the aim of capturing PE teachers' perceptions about their own assessment practices on a larger scale**.** A cross-sectional design gives a snapshot of a particular population at one point in time (Cohen, Manion and Morrison, 2018). It can be said to be a good way to acquire information of people of different ages at the same time, and is generally an inexpensive and quick way to do so (Salkind, 2010). To this end, a survey was developed,

piloted and distributed to secondary school PE teachers. The aim of the survey was to explore the nature and range of secondary school PE teachers' current assessment practices, their perceptions and understanding of the purpose and application of assessment, their prior initial teacher education and professional development opportunities with a focus on assessment, and whether teachers see assessment as being conducive to learning. The survey was developed and piloted once results from phase one were analysed. Piloting of the surveys was done with seven PE teachers and explained further in the following sections.

The online survey was distributed to all the secondary PE teachers teaching in State and church schools. Malta has a small population of around four hundred and thirty-two thousand people with the number of Maltese PE teaching in state schools amounting to eighty, and those teaching in church schools to forty seven. The surveys were intended to provide valuable information about a particular population (Frippiat and Marquis, 2010), while also providing descriptions of a numerical nature, attitudes and opinions about the same population (Creswell, 2003).

The survey was distributed to all the teachers via email through the Education Officer (EO) for the state schools, and the Head of Department (HoD) for the church schools. It was established that all the teachers have internet access and are in possession of a school email account. With this in mind, and to ensure that the survey was conducted in the most cost-effective way possible, an online survey was used to reach all the PE teachers on both of the Maltese islands (Van Selm and Janowski, 2006).

Various studies showed that response to online surveys is getting lower, with some reaching as low as 10 percent (Van Mol, 2017). Cohen et al. (2018, p. 501) go as far as to state

'Be satisfied if you receive a 50 percent response to the survey', adding that the researcher in education is to expect a much lower response rate.

However, at 43, the initial response rate was not as high as anticipated. It was therefore decided that to access as many PE teachers as possible, other strategies had to be used. Teachers in Malta have to attend a yearly inset course usually held in July. Seeing this as an opportunity for increasing response rate, a number of surveys were also printed and given out at the start of the course, using the drop-off/ pick-up method. This personal interaction results in an increase of the response rate (Allred and Ross-Davis, 2011), and the final number of PE teachers completing the survey was 90.

Nutly (2008) adapted a formula and set a ten percent sampling error and suggested an accepted eighty percent confidence level. A table was created to serve as a guide to what can be considered an acceptable response rate. She came up with 2 different 'required response rates' (Nutly, 2008, p.310). Type one is where you have 'liberal conditions', a ten percent sampling error and eighty percent confidence levels. Type two is where more 'stringent conditions' are reached (ibid.). In this scenario a three percent sampling error and a ninety five percent confidence level are achieved. Since the survey was sent out to the entire population of Secondary State and Church School Physical Education teachers (n=127), a minimum of 17 responses were needed for liberal conditions and 101 responses for the stringent conditions.

The online survey was completed by a total of ninety teachers, out of the one hundred and twenty seven teachers in both state and church schools (71%). Of the 90 teachers that agreed to participate in the study, n=69 had a complete demographic data and provided

responses to all items, and were therefore eligible for inclusion in this study (overall 54% response rate).

Although the response rate was not as high as 85%, it was still a considerable response. Splitting the percentage response between church and state-owned schools, a clear split in the responses could be noted. A total of fifty four teachers from state schools and fifteen teachers from church schools responded to the survey, which translated into a 67.5% response from the state teachers, and a 32% response rate for church schools. A summary of the sample's characteristics is shown in Table 2.

Table 3: Characteristics of Survey Participants

Characteristics	Percentage	Number
Gender		
Male	67%	46
Female	33%	23
Other	0%	0
School Type		
State	78.%	54
Church	22%	15
Level of Education		
Master's Degree	19%	13
Bachelor's Degree in Physical Education	68%	47
Bachelor's Degree in another subject	2%	1
	2%	1

BSc in Sport	9%	6
Diploma	2%	1
Other		
Teaching Experience		
1-6	39%	27
7-17	32%	22
14-20	15%	10
23-28	15%	10
Teaching Status		
Regular Teacher	88%	61
Part time Teacher	2%	1
Supply	10%	7

Key demographic characteristics for the teachers are shown in the above table, where it can be seen that 33% of the teachers were female, whilst 67% were male; the average age of the respondents was thirty five, with the youngest participant being eighteen and the eldest being sixty two years old. It is evident the majority of teachers have a bachelor's degree as their highest level of education (68%), followed by a Master's Degree (18%). More than half of the respondents have less than seventeen years of teaching experience, with nearly forty percent of the respondents being in the range of having one to six years of experience in the field. The vast majority of participants (over 88%) were regular teachers, followed by supply teachers and one part-time teacher.

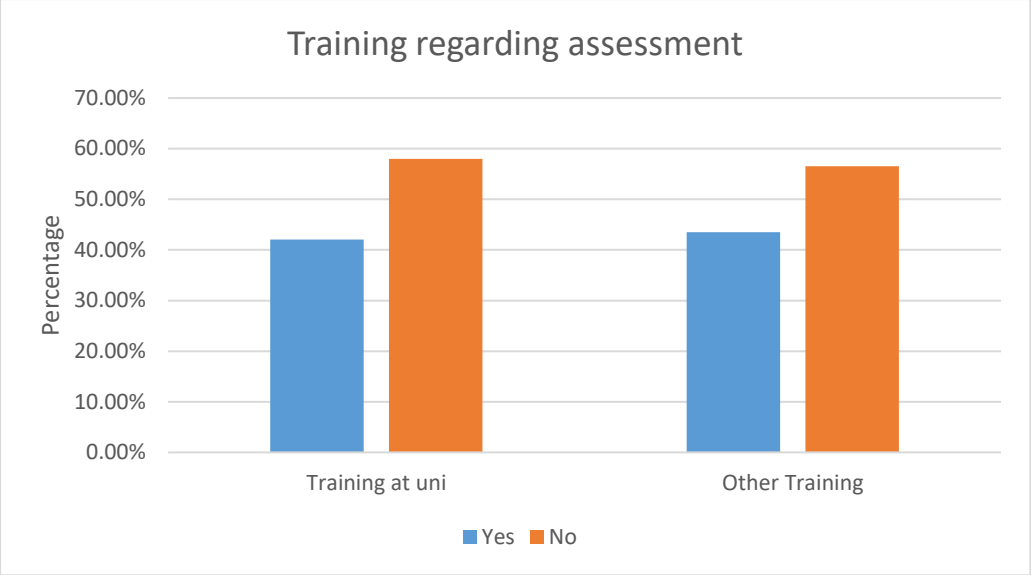

Figure 1 Teacher training on assessment

When asked on whether they received any training at university regarding assessment in general PE and how to conduct it, the majority of teachers (over 50%) stated that no training was given, with less than 40% of teachers saying that training was given. This was shown in the interview stage where some teachers could confidently say that no training was given at university or otherwise, while some other weren't sure but stated that they most probably were, directly or indirectly.

When going through the university units on assessment, it resulted that only one unit MSY3122 that only tackles assessment in SEC PE, which is for those students who choose PE as an option for their national exams.

3.2 Data Collection tools

As previously noted, three different data collection tools were employed in the two research phases. These will now be explained and justified in turn. In the first phase of the study, interviews and observations were employed. These served the purpose of collecting qualitative data on the teachers' approaches, practices, thoughts and ideas on assessment, as well as to formulate questions for the survey that would further delve into the assessment approach. The observations served as a means of triangulation on what was derived from the interviews. Triangulation would help viewing practices from other perspectives (Denscombe, 2007), so as to be able to get a better understanding of what was stated in the interviews.

3.2.1 Interviews

Interviewing is a widely accepted method of data collection in educational research (Coleman, 2012). It has been described as more than 'just a conversation' (Denscombe, 2007, p. 173); a tool for the collection of data which can be flexible enough to allow various multi-sensory channels of communication to be used, including 'verbal, non-verbal, seen, spoken and heard' (Cohen, Manion and Morrison, 2018, p. 409). This allows researchers to delve into what participants understand, believe or think while leaving room for 'spontaneity'; the interviewees' freedom to talk and express themselves in any manner (ibid., p. 506).

In this study the researcher made sure that the interview was conducted in an environment deemed comfortable by the interviewee. Both Cohen, et al. (2018) and Denscombe (2007) agree that during an interview the researcher can delve more into what the participant is stating and search for deeper underlying causes, explanations or other related matters to gain deeper insights. Interviews can be said to be tools that encourage

both the researcher and the interviewee to engage in a discussion on the way they view things on a particular topic (Cohen, Manion and Morrison, 2018). It is also a tool which enables the interviewer to gain access to the participants' 'feelings, emotions and experiences' (Descombe, 2007, p. 174).

Specifically, one-to-one, face-to-face, semi-structured interviewing, one of the most common forms of interviewing was employed (Denscombe, 2007). This permitted the interviewer to ensure a list of relevant and important matters to be discussed, whilst at the same time giving participants the freedom to raise other related issues (Denscombe, 2007). Having a set of pre-determined questions was important to ensure that the interviews would provide the desired data on the important matters in relation to the research questions, thus enhancing the study's reliability (Puzanova, Larlina and Zakharova, 2018).

It is important to note that the interviewing process is not without its limitations. These might include the interviewer asking leading questions which consciously or unconsciously affect how the interviewee might respond to important questions (ibid.). Being aware of this potential threat to the study's validity, a detailed interview protocol was developed and reviewed by the project supervisor and later piloted to ensure that the questions asked were generic and did not lead to specific types of responses (Appendix 1). Semi-structured interviewing is also time consuming, and this fact influenced the decision on how many teachers would be involved in the study in order to ensure that sufficient time was allocated to collection and analysing of the data.

The interview protocol

The interview was split into four main parts. The first part sought to collect information about each teacher; their background, years in the profession, perceptions on the PE curriculum, and preference of topics. Gaining insight into the participants' views on assessment was the aim of the second theme of the interview. Questions and prompts to explore their views and perceptions on the purpose and importance of assessment more broadly and assessment in the context of physical education were asked. Evidence on teachers' perceptions on specific assessment strategies they employ, and their views on the meaning, importance and impact of these practices was also collected. The third theme focused on understanding teachers' views on the content, nature and quality of initial teacher education and any follow-up professional development experienced they had on assessment. Teaching styles was the main focus of the fourth theme in the interview protocol. As the Physical Education syllabus is focused on four main groups as mentioned in the previous chapter, participants were probed on what teaching styles they use when teaching specific topics and how they assess thee topics. Probing questions, for example, of how the lessons are conducted, assured reliability and served as a cross-check to the methods the participants mentioned earlier. This theme was also aimed at trying to see whether there is a relationship between the teaching style and the method of assessment used later. The last theme focused on the teachers' perceptions of their own effectiveness as far as assessment practices were concerned, and that is, to what extent they believed that what they did in relation to assessment had a direct impact on student learning and overall school performance.

Open-ended questions were used in the interviews. According to Cohen, Manion and Morrison (2018) these present a number of advantages. Primarily, they allow the researcher to investigate and go into more detail wherever needed. This also helps in clarifying any potential misunderstandings or misconceptions. Building a relationship with the interviewee

and inspiring cooperation are also two other advantages gained by using open-ended questions. A funneling approach was used in the interview protocol so as to focus on particular aspects in the field. The first section in the protocol was made up of generic questions aimed at gathering information about the respondent, while the other sections delved into the areas revolving around assessment. During the interview process, probes and prompts were used. According to Cohen, Manion and Morrison (2018), these are used when the interviewee does not fully understand a question or misunderstands what is being asked. Probes as prepared in the interview protocol, were aimed at helping the interviewee to elaborate or expand upon particular issues that might arise in the interview. Some of them were anticipated probes, and prepared as part of the interview protocol (Beatty and Willis, 2007).

Piloting

The interview protocol was piloted with two secondary school PE teachers. Piloting this interview gave the researcher the opportunity to refine the interviewing skills, like sensitivity to feelings, adaptation, probing and using checks (Denscombe, 2007). Probing regarding teaching style was done in order to make sure which style was used, as there seemed to be a discrepancy between what the teacher named as a teaching style and the actual style used in the example given by the same teacher. Doing the pilot interviews also gave the researcher

an opportunity to clear out any misconceptions or queries that arose from the protocol (Turner, 2010). After the two pilot interviews the questions on training at university changed from a question directly on specific units at university to a generic one where the interviewee could mention anything done in relation to assessment, helping to retain the flow of the interview.

3.2.2 Observations

It is widely accepted that observation is a research tool that can give researchers direct access to the phenomenon under investigation, rather than rely on secondary sources; and to collect evidence that relies on the direct eye witness of the researcher (Denscombe, 2007). It has been argued that this direct access to the phenomenon under investigation can uniquely provide the means to produce data that is more 'valid and authentic' than other methods (Cohen, Manion and Morrison, 2018, p. 542). There are, however, certain conditions that maximise the power of this data collection tool. As Denscombe (2007) explains: *direct observations* permit the researchers to be present during key events and to collect evidence in relation to the questions investigated. Whether or not the research project takes place, the particular event would still happen, ensuring a *natural setting* for the observation.

When observing these natural phenomena, there is of course the crucial matter of researchers' subjectivity and the issue of perception as Denscombe (2007) explains. Personal factors and interpretations can lead to the data collected being unreliable. An important aspect in ensuring the trustworthiness of the data collected via observations is the extent to which the same data would be collected by two different researchers/observers. Although

the competence of the observer plays an important role, perception plays a crucial part in the recording of observations. Dencsombe (2007) describes the process by which the observer organizes and selects certain stimuli, and emphasizes that the variation in observations depend on the following:

- *Familiarity*: A researcher conducting observations will most probably observe and see what he is accustomed to seeing (p. 208). In relation to the present study, if the student-researcher, a PE teacher himself, focuses on particular elements in his teaching, then there will be a natural tendency to observe the session of other practitioners in relation to the lack and/or presence of that particular element.
- *Past experiences*: Past experiences can hinder the observer too. If the observer has been taught the particular subject in his past, there may be a focus on desirable stimuli, omitting others that are considered as undesirable to the observer.
- *Observer's presence*: that the naturalness of the setting (Denscombe, 2007, p.215) could be tampered by the researcher's presence.

Observation Schedules

In order to avoid any of the points affecting an observation mentioned above, an observation schedule was created. This was taken to the observation sites and completed by the researcher to make sure that important details are observed. Denscombe (2007) states that

an observation schedule is a tool which is employed in a bid to eliminate any errors in recording due to different researcher interpretations. He further describes the schedule as being purposefully designed to eliminate or minimise as much as possible any variations that would arise from different observers. Schedules allow observers to be alert to the same activities and to look out for the same things. They also allow data to be recorded in a systemic and thorough manner, producing data which is constant and consistent even if the observers change.

The observation schedule in this study contained a list of items similar to a checklist (Denscombe, 2007). This allowed for systemic observations, checking the frequency of how many times an item occurs, or the duration of it. It also contained a separate space for ethnographic field notes regarding assessment, lesson structure, and any related student or teacher interaction.

It is suggested that the schedule is built and based on the literature review as this often provides the researcher with aspects that are worthy of observing. This will help the schedule be more focused and relevant. The observation schedule used in this study can be found in Appendix 2. The first section contained *frequency of events* where the researcher counted the number of times an event occurred during the observation sessions. In this part the main data collected revolved around the frequency of: (i) teacher explanation, (ii) students actively engaged, (iii) on-task student communication (the number of times that students communicated and worked together on a task), (iv) teacher feedback (the number of times that the teacher gave feedback to students), and (v) student feedback (the number of times that students gave feedback to each other).

The second type of data collection revolved around taking open-ended field notes on the lesson structure and progression. The researcher took notes of teachers' instructions, students' responses and task execution.

Systemic observation has been used in this research. It required the researcher to spend period of around forty five minute in the field. Furthermore, this required the observer to 'fade' into the background becoming invisible to the participants (Denscombe, 2007, p.215).

3.2.3 Teacher survey on assessment

The survey itself had four distinct parts designed to achieve the said two purposes. Most questions were derived from the data gathered in the interviews and observations.

Being a teacher, the researcher had already formed some hypotheses regarding what was happening in schools regarding assessment. Thus, he feared that he might have pre-existing prejudices which might lead him to present a rather subjective picture. This was another reason why a survey was carried out. McNeill and Chapman (2005) mention how surveys have the capability of producing data which is objective, and, moreover, confirmed by the participants taking part.

This research aimed to collect as much data as possible from all the teachers teaching Physical Education in both state and church schools. For this purpose, an online survey was created and sent to all the PE teachers (n=127). This is made up of forty-seven teachers teaching in church schools, and eighty teachers teaching in state schools.

A list of email addresses and names of the teachers teaching in a secondary school setting in state schools could not be obtained from the relevant EO. Neither could this data for teachers in church schools be obtained from the relevant HoD due to data protection regulations. Instead, permission was obtained both from the Directorate for Quality Standards in Education, which falls under the Maltese Ministry for Education, and from the Director for Educational Services at the Secretariat for Catholic Education. Both the EO and the HoD were then sent a copy of the survey link together with an introductory letter, which in turn was forwarded to all PE teachers (n=127).

However, there are some minor flaws or weaknesses that can be attributed to the use of surveys which should be kept in mind. Pre-coded questions can be biased towards the researcher's beliefs (Denscombe, 2007), although this should have been balanced out through the interview process. Denscombe (2007) also posits that surveys may limit the capability of the researcher to cross check for 'truthfulness of the answers given by the respondents' (p.171). He blames this on the fact that since the surveys are done at a distance, the researcher cannot tap into the expressions and any other clues that may crop up in an interview. Furthermore, the survey may limit further probing by the researcher into any particular responses. In this study, the researcher sought to minimize this by providing the respondents with space to elaborate on yes or no responses.

Structure of the Survey

General information

Part one was intended to gather general data about the participants. Information on age, gender, teaching experience and opinion on the domain importance in Physical Education was

sought. The question on domain importance was derived from the fact that a number of credits in teacher training at the University of Malta revolve around the importance of teaching the four domains as set out by Mosston and Ashworth (2002), namely the physical, cognitive, social and affective domains. Some of the questions used were just a multiple choice or simple data entry. A four point Likert scale ranging from very important to not important at all was employed for the question regarding domain importance given by teachers. This type of scale is most probably the most popular response scale (Chyung, et al., 2017). The other questions were designed to have a selection from a drop down item menu or just a fill in box, as the questions were just aimed at obtaining facts on age, gender and teaching experience.

Training in assessment

The second part of the survey was structured to obtain information regarding any training the participant may have had regarding assessment, both at university and at a personal development level. The questions were yes-no questions aimed at obtaining data regarding training in assessment at university level or beyond. Text boxes were provided to gain further insight to the type of training received if a yes option was chosen. Details sought were on duration and main title of the training given. A four point Likert scale ranging from very important to not important at all was employed for the question regarding the effectiveness of the courses attended or units done at university as part of teacher training.

Type of assessment used

The third part of the survey sought to gather facts regarding assessment used by teachers and what is usually assessed during the year. Multiple choice questions and four point Likert scale questions were used to obtain information regarding syllabi followed, assessment schemes

and timing of assessment. On the other hand, six point Likert scale questions were used in the questions with related to what is assessed in the different areas of the curriculum in order to obtain an accurate measure of the frequency of the different domains assessed, while a five point Likert scale ranging from never to always was used to measure the types of assessment methods used in different areas across the curriculum. Having a five point psychometric scale creates anchors in this measurement method, with the middle anchor many times being the neutral one (Chyung, et al., 2017).

Student learning

In the last part of the survey, the effect of assessment on students learning was tackled. This was aimed to get a view of the teachers' thoughts on the efficacy and impact that assessment has on students. In this final part of the survey multiple choice questions were used alongside text boxes for the teachers to explain in depth the reason for their response.

Piloting

It was vital that the survey was as reliable as possible before the actual distribution of the surveys. The survey was piloted and tested three times with seven different teachers teaching at secondary level. This was done so as to ensure that any issues ranging from formatting to clarity of questions posed or any ambiguities in the questions Cohen, Manion and Morrison, 2018) were acted upon before the actual distribution. This piloting phase is vital in any research as it is a 'necessary but insufficient condition for validity in research; it is a necessary precondition of validity' (ibid., p. 247). After the first pilot study a whole section was

eliminated as it made the survey significantly longer with no significant data for the purpose of this study being collected. It was also decided that the option 'other' be added to certain questions like 'type of assessment scheme' to allow for more precise data to be gathered.

3.3 Ethics

Preceding the start of the study, ethical approval was obtained from the University of Birmingham Ethics Committee. Permission was also obtained both from the Directorate for Quality Standards in Education, which falls under the Maltese Ministry for Education and from the Director for Educational Services at the Secretariat for Catholic Education. These were needed to gain access to both state and church schools respectively. Once the said approvals were obtained, the respective heads of schools had to be contacted for permission to conduct research within their school. Formal consent forms were given to the interview participants with a letter explaining the nature and purpose of their participation. A detailed letter explaining the nature and purpose of their participation was sent with the link to the teacher survey where the participants were advised, they gave their informed consent to participate in the study by filling in the survey. All participants also had the right to withdraw from the study (before May 2018 for the interviews and observations, and before August 2018 for the surveys) by contacting the main investigator (Mr Karl Cortis). Furthermore, the parents had to be informed regarding the observations being conducted during their children's lessons. This was done with an email sent to the parents, and in one case, a letter of information was handed out the parents.

3.4 Analysis of Data

3.4.1 Analysis of interviews and observations

Qualitative data was analysed using thematic content analysis, where the aim was to find common patterns across the interviews and observations (Hsieh and Shannon, 2005). Interviews were analysed in a manual manner, since the data set was considered too small to make use of software options such as NVivo. Data was first compared according to the question asked so as to check for any common elements, discrepancies and viewpoints. Thematic coding was then used to interpret the transcriptions. Marshall and Rossman's (2016) six-stage process was used to identify any themes emerging from the data. After the transcription of the interview, the transcripts were read several times. The data was categorised into themes, according to the questions asked, and then coded. Any emerging themes were read and checked by the researcher to make sure that they correspond to the research questions. Alternative explanations on the emerging data was also sought. Any additional themes which emerged from the surveys were later added to the existent themes.

Stage One

After the interview, the transcription was immediately carried out in order to organise and make sense of what the participants had said. Working with a transcriber and sound recording in Mp3 format was helpful in listening to the participants' narratives for as many times as was deemed necessary. This was important in order to get truly attuned to the participants' responses and to become sensitive to the interviewees' perspectives.

Stage Two

This second step involved the merging of multiple data in codes. Through this, main points in relation to the research questions were selected and labelled.

Stage Three

Here codes were devised and categorised. Data was labelled in codes so as to be able to create common categories to correspond to the main research questions. Codes like 'time constraints', and 'lack of support' were grouped in a common category named 'problems faced by educators'.

Stage Four

In this step the categories were grouped under the umbrella of the research questions, depending on which question they were related to. Although usually an inductive process, since the research questions were already in place, with the categories created to answer them, this process was in fact deductive in nature.

Stage Five

A close analysis was done on the data gathered, bringing together any issues regarding the key findings regarding the research questions. This is presented in the results chapter.

Stage Six

Analysing the data collected helped identify the key points, contradictions and explanations for the findings in the study. This is presented in the discussion chapter.

3.4.2 Analysis of Surveys

The data gathered by the lime survey software was analysed using SPSS version 25 (IBM Statistics). Descriptive tests were run on the data to have a clear idea of the most common groups and concepts in the study. Friedman tests were carried out to compare mean rating scores provided for a number of statements all measured on the same Likert scales. In this study the use of bar charts was the mode chosen for data representation. The data was first split into raw percentages. After this cross tabulation was used to split the percentages into more focused groups. The data was then filtered and focused on one group at a time. Chi squared tests were also carried out on some of the data. Statistics regarding school type, years of experience and training were compiled to give a clear overview of the participants involved.

3.5 Summary

This chapter has shown the philosophy behind the methods chosen, their design, all the methods used for the collection of data, and the analysis of the data collected. It also aimed to demonstrate the logic and reasoning behind the techniques used to respond to the following research questions:

1. What are PE teachers' understandings of the notion of assessment and assessment policies in Malta?

2. What assessment practices / approaches do PE teachers employ and why?

3. What barriers do PE teachers encounter that hinder their assessment practices?

Chapter 4. Results

Throughout this chapter, the data which was collected through the two phases is being presented. The results are split according to the research questions where results of both the qualitative and the quantitative methods are joined together to create a more holistic picture of the data gathered.

In this study, teachers seemed to share very similar ideas on the multi-dimensional nature of physical education – that is, they identified four domains in terms of what the focus and purpose of PE should be. This seemed to transpose onto assessment as overall this was understood by most teachers to be used to assess beyond the physical dimension. It was also evident that the extent of the teacher's experience (i.e., years in the profession) had a direct impact in the mode of assessment of certain domains. Some concerns on the effectiveness of the current assessment policies and practices were highlighted, including questions about the nature and quality of teacher preparation to be able to assess in effective and meaningful way. Of note, even though teachers reported following similar syllabi, and overall seemed appeared to share common perceptions about the importance and mode of assessment, including its potential to affect student motivation, however, the way they conduct assessment seems to vary. Finally, teachers also expressed some concern regarding the current focus of assessments on performance rather than progress vis-à-vis the national policies.

4.1 Teachers' interpretation of assessment

Most teachers interviewed (n=6), when asked to provide their own definition or interpretation of assessment, they talked about assessment as incorporating primarily a summative process, held at the end of a unit or topic. For these teachers, summative assessment was in turn a significant element of their work for accountability purposes; a

requirement set by the school so that parents are informed about their child's progress but also the school has information on what teachers do. In this case, assessment 'serves primarily as a summary of what you have done during the year' (T7, F). All teachers explained how they were expected to mark formally at least twice a year but two teachers did not see value in that process:

> **'For me [this type of summative/formal] assessment is a waste of time. If you are doing assessment after each topic, you are wasting a week from each session. In fact, I only do one assessment every term. I make sure to alternate assessments [peer assessment, observations, and self-assessment] so that all areas [physical, cognitive, social and affective] are assessed till the end of the year. In that way I manage to target all the students.'**
> **(T7, F)**

When asked whether they felt that summative assessment in Physical Education was important, mixed opinions came across:

> ***'Yes and no... So assessment is important for the school, but one has to keep in mind that not all students are good in everything. If assessment is on one topic which may be the weakest link of the student it will not show the truth on the student. She can get a C in PE because assessment was on the topic she was weakest in. In fact I have started splitting assessment so that the parents know what has been covered and on what criteria the mark is based on.'***
> ***(T7, F)***

With another respondent emphasising the teacher's attitude and the way it is presented to the students:

> ***'It is important but it depends on the context. And moreover, the attitude the teacher has towards assessment. If a teacher just does the assessment to say this student got an 80 on 100, or a pass or fail, the aim of assessment is really lost. If you do the assessment to show the student where he stands and to see how you are going to manage to bring him to that point, then yes, assessment is useful.'***
> ***(T6, M)***

Some respondents (n=5) had very strong views against any type of summative assessment in the context of PE in schools. For example, the quote below illustrates these teachers' concerns regarding the negative affect of summative assessment – when implemented for bureaucratic purposes only and without any consideration on how to support students to progress - on student motivation. This teacher also acknowledged that summative assessment was also meaningless for some students too who did not see value in PE.

Although no respondents chose the option stating that assessment practices effect students' motivation in a negative way, and the majority chose that they may affect motivation in a positive way. Some teachers chose to check the option that assessment may sometimes effect students' motivation in a negative manner.

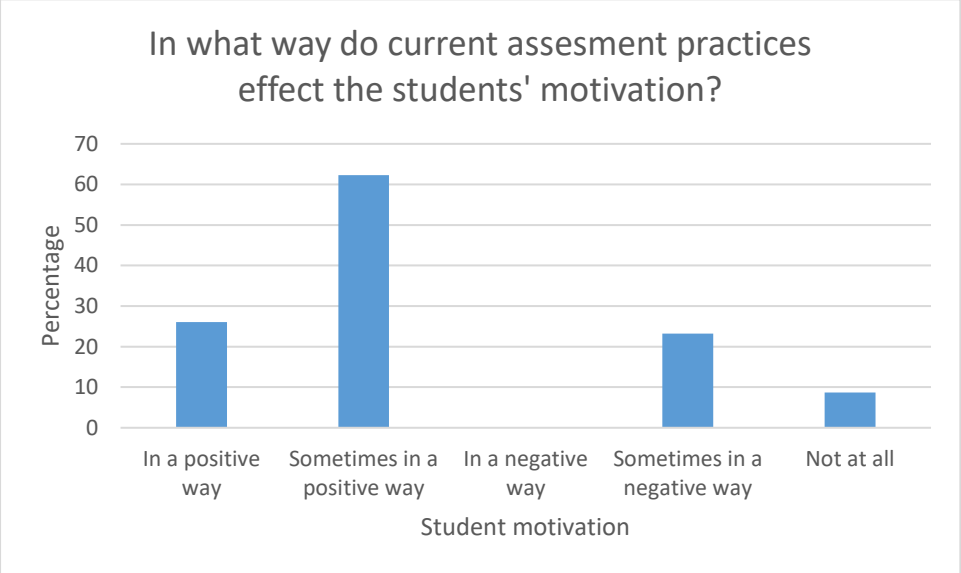

Figure 2 Way in which assessment effects student motivation

Other teachers (n=3) emphasised the importance of assessment in supporting pupils to understand where they are in their learning. So, for these teachers, the information obtained

through the assessment process should be used so that the students understand what they have achieved:

> ***'Assessment should tell you where the student's stands, but more than that It should help to give a way on how to help the student reach the objectives.'***
>
> ***(T6, M)***

> ***'Assessment is when you [meaning the students] know how you are doing in a subject. Whether you are improving and arriving to another point.'***
>
> ***(T3, F)***

One teacher underlined that what is assessed should not be constrained on what pupils can do with their bodies but also reflect their level of understanding:

> ***'For me assessment is watching a student progress in a particular skill, let say correct running technique, constantly and in good parameters. By parameters I mean that for example in term two, the teacher can understand that 'Joe' has managed to arrive to where the teacher expected in terms of achievements. The achievements can be simple cognitive achievements, like understanding, but not performing.'***
> ***(T5, M)***

> ***When observing T5 there was a clear use of comparison to what the assessment sheet that the teacher created for the topic. The emphasis was clearly on technique and not on time performance! There was a clear focus on understanding the process behind the technique and also the ability to execute it well.***

One teacher argued that this valuable information on whether objectives have been met or not should be important information teachers rely upon to inform their own teaching – and to assessment the effectiveness of their teaching (i.e. what impact their work has had):

> ***'Assessment is a tool to get feedback, both on yourself as a teacher teaching the subject and for the student who is trying to do or learn a particular task.'***
> ***(T6, M)***

'I see assessment as feedback, a tool to get feedback, both on yourself as a teacher teaching the subject and for the student who is trying to do or learn a particular task.'
(T2, M)

With the use of effective assessment, teachers can consequently make decision on what and how to improve their provision:

'In my opinion if you are doing something and you are doing it wrong, if there is no assessment you won't realise, and you will continue doing the same mistakes. If they [students] get some form of assessment and they find that something is wrong, they can start tweaking on what they were doing so that they improve. Feedback is highly important as so to improve.'

(T2, M)

A variety of opinions on what should be assessed in PE were shared by the teachers interviewed, ranging from assessing students' level of participation and fitness to including a clear strategy to assess not just the physical domain but other potential outcomes. It is interesting to see teachers' responses on the important matter of what should be assessed in relation to their responses on the overall purpose of PE.

All teachers interviewed believed that one of the most important purposes of PE was to promote lifelong engagement in physical activity, and this was justified by drawing upon concerning figures of childhood obesity:

'I believe we should have PE every day and from a young age. This is so that it becomes part of our culture. We have a very big problem of obesity in this country, the highest in Europe. If we get them [students] moving and active at a young age, it is the way forward to fighting obesity….We need to work on the health of the country. I believe that exercise gives you discipline, something they need for life.'
(T4, F)

Other teachers were concerned about the lack of physical competence amongst young children and believed that in this context of rising obesity and restricted physical competence, the role of PE was significant:

> ***'It is part of the holistic development of the students. I've only been teaching for 3 years but I can already observe that the students are lacking skills that they should already know. Like fundamental and basic skills.'***
>
> ***(T8, F)***

So, to support students develop these skills and to promote an active lifestyle, these teachers believed that PE provision needs to be fun, engaging, and to offer pupils the opportunity to try out a wide range of sports. The quote below illustrates the majority of teachers' views about the importance of exposure to a variety of sport activities so that pupils find something they enjoy and are active 'for life':

> ***'The aim [of Physical Education] is to have the students move but also to offer them all the sports. Sometimes a student might say, "I really like this sport, I want to try it out". I think that is the main aim of PE. That they have something to relate to so after school they can vent off there.'***
> ***(T7, F)***

So, the overall perception was that assessment in schools was promoted as a bureaucratic exercise, so that schools understand what teachers teach and parents have an idea on where their students are. However, it was evident that in many cases, teachers were not expected to offer a level of detail or specificity to properly inform parents about where their child was and in what aspect of PE attainment targets achieved well or not. However, it was also evident, that beyond this narrow interpretation and expectations placed upon them, most teachers believed that assessment, if done well, is a necessary element of teaching and learning.

4.2 Assessment Practices

Around half of the survey respondents reported choosing to adapt the national set assessment and make the assessment criteria more school based. Just under 50% of the teachers chose to keep the national set assessment. The teacher below explained why, in her school, they adapted the national set of criteria to improve their assessments procedures in her school:

> ***'We the teachers decide. We use the national syllabus as a guideline, and to keep in line with what other schools are doing but I think that our students have more fundamental skills when compared to other schools, so we adapt according to our students, sometimes even to the class. Sometimes we increase the standard they are assessed against sometimes we decrease it, depending on the class and year group.' (T8, F)***

Although the teachers interviewed criticised the requirement to provide summative assessment as a bureaucratic information that is not meaningful for teachers, students and parents, they overall believed that assessment, if done well and for the right educational purposes, it can give important information to both teachers and their students (in terms of learning progression). When enquired to offer details on the ways they assess students, it became apparent that the formal summative assessment requirements were done normally at the end of the topic. Results from the teacher survey also suggest that the vast majority of teachers (nearly 60%) stated that assessment was usually done at the end of a topic. But they also claimed that they never used assessment at the beginning of a topic, which suggests that they rarely had a baseline against which to report progress. Formal assessment periods at the end of the topic seemed to be the norm.

It also became apparent that beyond the summative assessment requirements, according to these teachers, a wide range of methods were also employed in their PE lessons. Specifically, some argued that it is not enough if the student receives a summative mark at the end of the topic with no formative information to understand how and why this mark was provided; and crucially what they need to do to improve further. So, overall, four teachers mentioned the importance of formative assessment or assessment for learning. For example, a teacher underlined how ongoing assessment can keep students focused and motivated:

> ***'If you have ongoing assessment, the student is given an opportunity where to improve which will increase his interest and he will feel more engaged, as the assessment will keep them improving step by step.'***
> ***(T2, M)***

When asked to provide concrete examples on how, when, where and why they use assessment for learning, the four teachers explained that the use of assessment for learning (AfL) was not always done in a conscious and systematic way:

> ***'I think I do use it [AFL] without knowing. For example, if I am doing assessment and I tell the student how to do the skill, you are seeing what the student has understood, what she knows and tell her how to do it. So it is still a type of assessment. It is more formative.'***
> ***(T7, F)***

Two respondents emphasised the importance of feedback to direct assessment and to promote learning:

> ***'As I explained before I see a gap between AFL and AOL. Personally, I give more value to assessment which is assessment for learning. If you are just doing assessment of learning and you just stop there, you give no feedback and the students did not work on what you said immediately, then the aim of assessment is lost. I believe that assessment should be for learning. After all we are there to promote learning over the four domains.'***
>
> ***(T6, M)***

'At the end you have to see what the student has learnt. But when you do this you are not in much of a position to help them improve after the assessment. That is why I give them feedback after the assessment is done, even in between trials maybe the student improves the second trial. I ask her to remember the previous lesson, maybe the students improves the mark.'
(T8, F)

A teacher delved into how he manages to do assessment for learning in the lessons. He mentioned the use of video analysis and how he involves students in the process as students observe their performances, with a focus on skill development and execution, assess themselves and give feedback to each other regarding the skill at hand:

'Assessment for learning is what I try to do. Saying "Hello, did you improve?" "Are you taking what we're doing seriously?" It's much deeper. I do use them a lot in PE, especially with video recording. When I record the boys doing for example a creative gymnastics routine, long jump, they can see themselves doing it.'
(T5, M)

The respondent continues to explain why AFL is so important.

'As a teacher you can see a whole different aspect of the assessment situation. For example, in the long jump, if you just write the students distance, you get the top, second, third etc., the students will already have forgotten about the whole thing. On the other hand if you take a video and have the students watch themselves, they start talking and discussing much bigger stuff like understanding the technique, the reason they are not getting a better distance. Why they did not do it well, maybe they were distracted. Through assessment AFL you can really understand how to go about teaching and improving. Using technology can help enhance this more.'
(T5, M)

This approach to assessment for learning was clearly evident in the lesson observation as shown below.

Field notes from observations

Teacher: T5, M

Students: 13

Year: 7

Topic: Gymnastics

[The teacher split the students into groups of four where they had to practice the gymnastics routine together. Each group was given a tablet on which to record the students doing the gymnastics skills learnt. Although students were occasionally off task in their interactions, overall, they were seen giving constructive feedback to each other, creating an atmosphere where teaching and learning was taking place. One student was showing the recording to the group explaining how the chin should be tucked into the chest when rolling, and how the space should be used better. His colleagues obliged, taking him seriously and asking whether the skill was better now. This teacher has created an atmosphere where a student, whatever the level, could learn and improve.]

This was a pristine example of using assessment as a part of learning and one of the three teachers to have actually done so in the lessons observed. Although teachers might be aware of assessment for learning and ways of actuating it, the actual implementation might not be being done.

The use of observation, peer and self-assessment was also explored in the survey involving the wider sample of PE teachers. Specifically, as shown in figure 2, across all subject areas, teachers reported using observation and feedback more frequently compared to self- and peer-assessment. In fact it ranked nearly 90% compared just below 50% of other types of assessment. But there was variation in how frequently the used these three forms of assessment per activity area as figures 3-6, show. The most common form of assessment was

observation and giving feedback. This was highest in Health Related Fitness (nearly 60%) and lowest in Outdoor Education (over 40%). This method was closely followed by observing and comparing against a set of criteria. Giving a mark on what was observed was a common practice used with nearly 50% in Team Games and Individual Activities and just over 40% in Health related fitness. Self-analysis and Partner analysis ranked as least favourite in all topics, with student self-analysis being slightly preferred over the partner skill analysis in all topics.

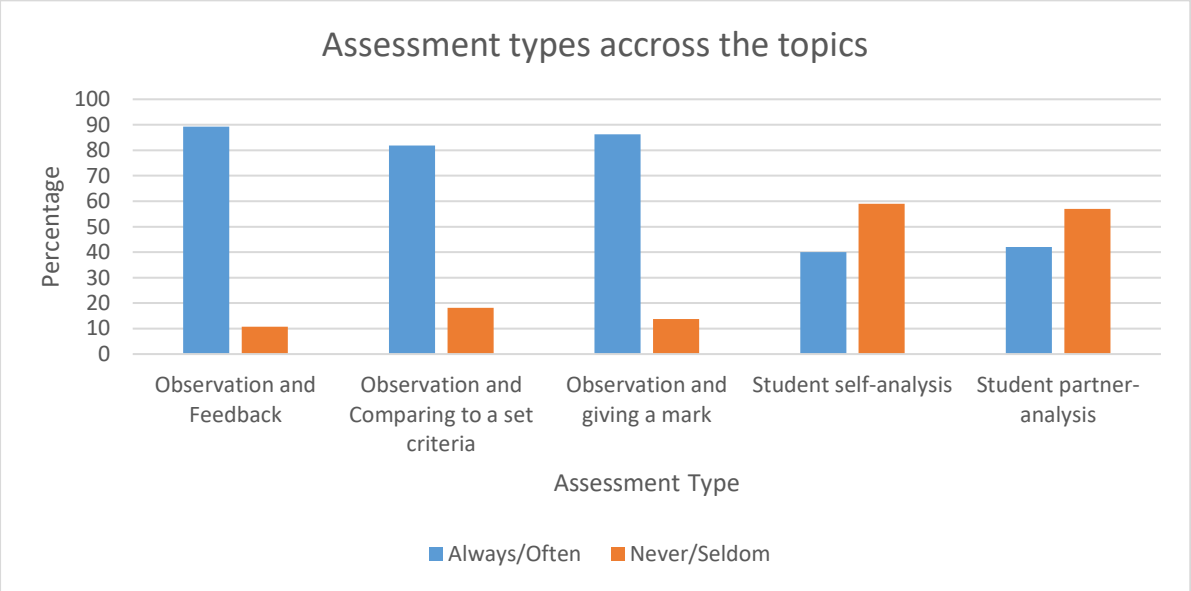

Figure 3 Assessment types across all areas

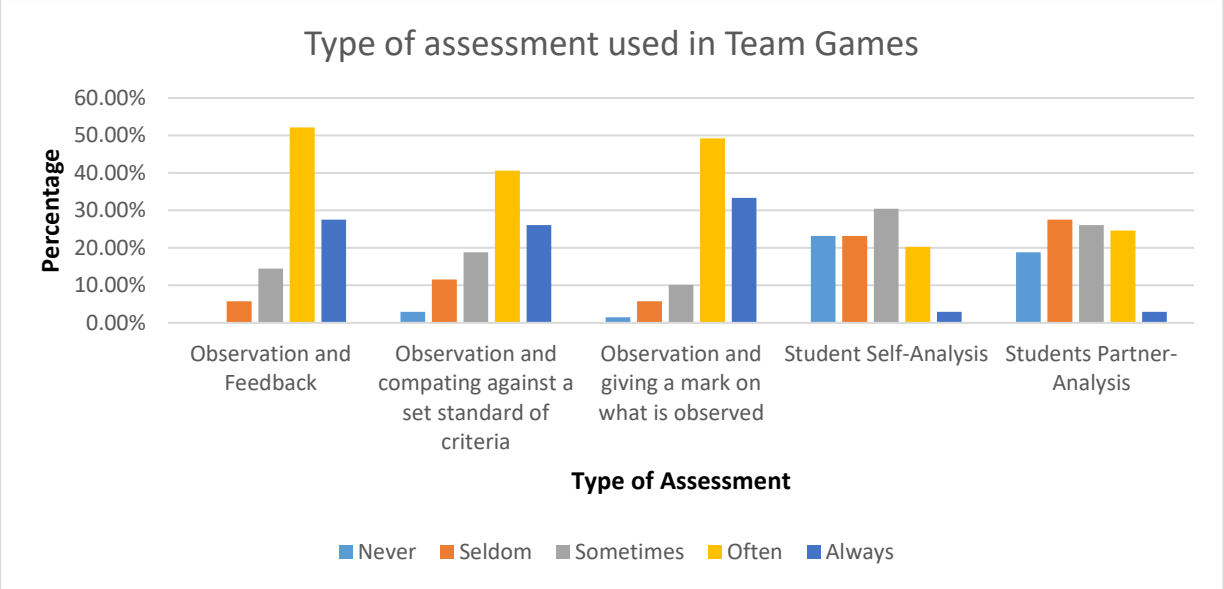

Figure 4 Assessment in Team Games

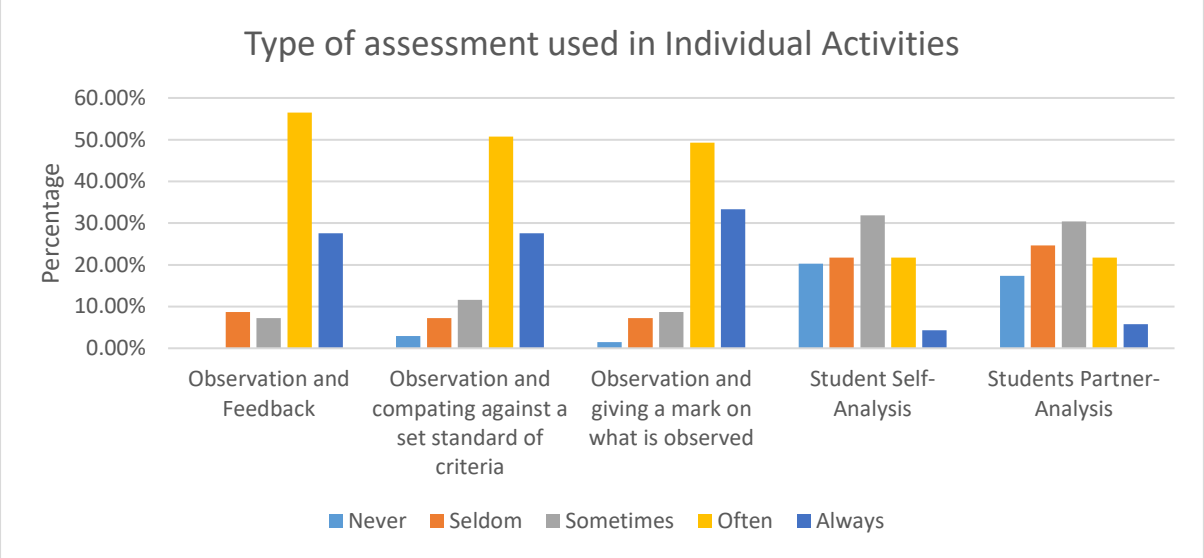

Figure 5 Assessment in Individual Activities

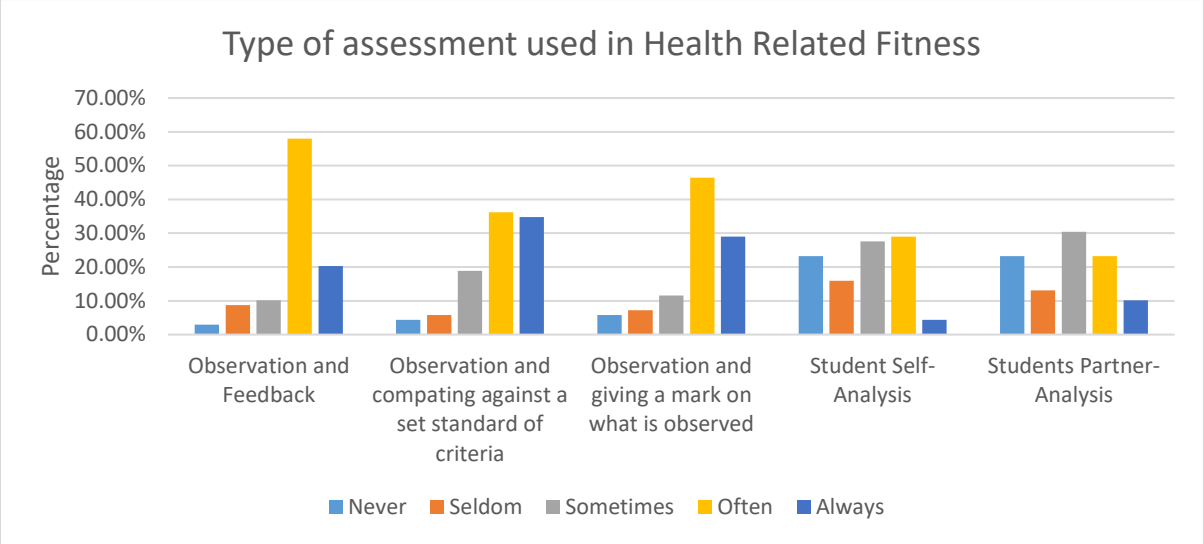

Figure 6 Assessment in HRF

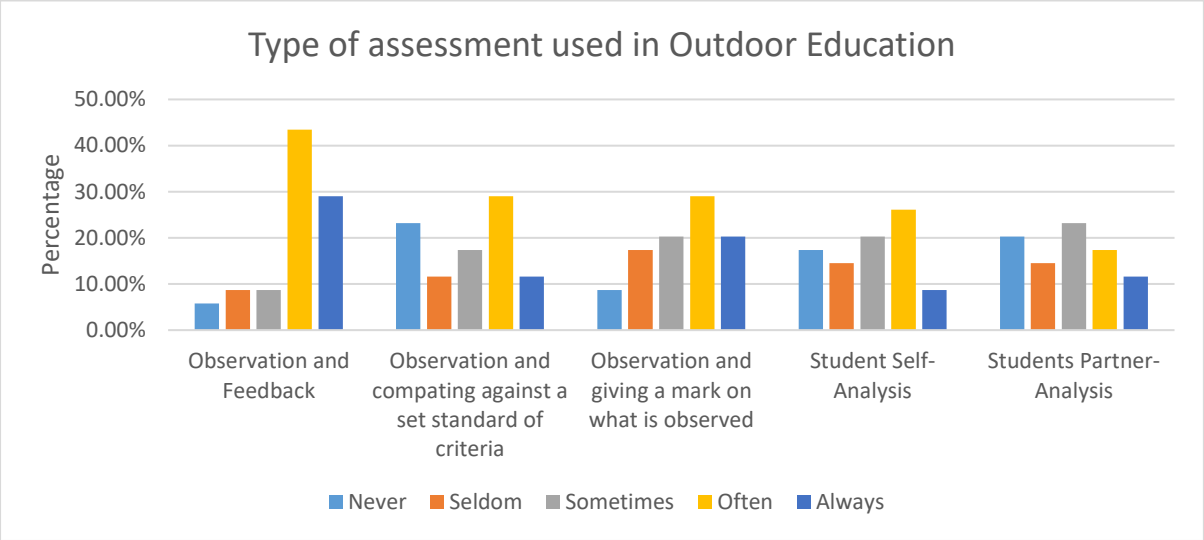

Figure 7 Assessment in OE

Some teachers commented about the importance and relevance of self- and peer assessment in the survey. They stated that it is very beneficial to the students' learning process.

I find partner analysis and self-analysis most useful. The student try their best to improve and compete against their own self. Usually more questions arise on how to do better.

The relevance and importance of self- and peer assessment was also discussed with three teachers during the interviews. As one teacher explained:

> ***'So I think that if in every group I place a leader or give them a responsibility as a team, and let them assess themselves, they still do everything and they are assessed at the same time.'***
> ***(T6, M)***

The same teacher argued that it was important to engage students in self- and peer-assessment in relation to assessment their own performance or the performance of others. This can enable them to think 'how you can improve the technique so that you have a higher success rate' (T6, M). The use of peer-assessment was evident during one PE lesson observation.

Field notes from observations

Teacher: T7, F

Students: 23

Year: 8

Topic: Badminton

The teacher has all the students grouped in groups of 3 or 4. After doing some rallies, the students were given hand-outs with the assessment criteria to be used. The teacher just explained how they have to tick in the box on whether the skill was achieved or not.

The students are actually assessing each other. The teacher is going around supporting the students and clarifying any difficulties they might have. The students are giving very accurate feedback to each other on the skills. Peer assessment and teaching at its best!

Responses from the survey reflect these teachers' views on the importance of achieving not only physical but also social and cognitive outcomes. However, as anticipated, the physical domain was ranked higher. Specifically, the vast majority of respondents (77%) reported that the physical domain was very important. This was followed by the social domain (56%) and cognitive (50%) domains where respondents marked the domains as very important

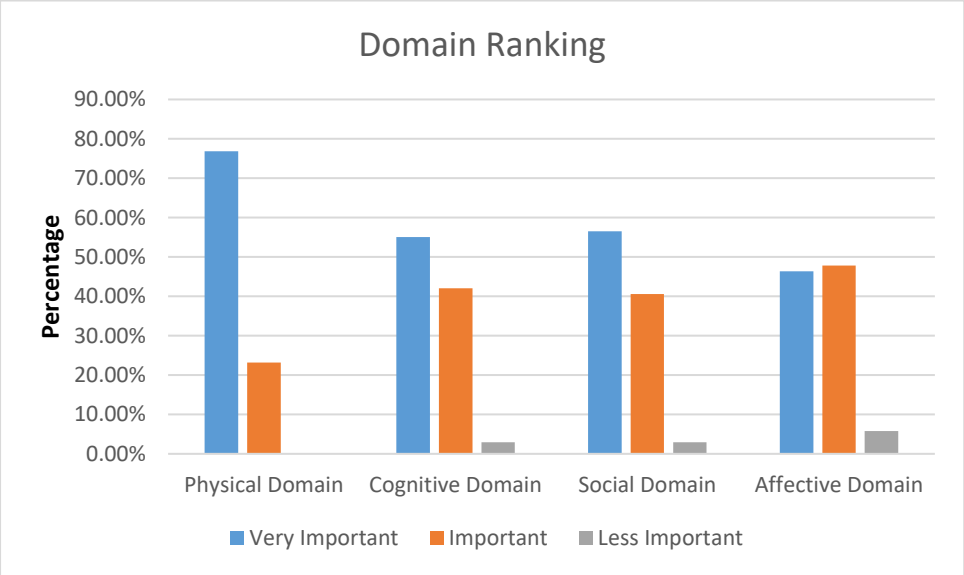

Figure 8 Domain Ranking

The mean rating score provided to the physical domain (3.77) is the largest. This was closely followed by the social domain (3.59), cognitive (3.55) and affective (3.42). This correlates to what the teachers stated in the interviews as to what the objectives of physical education are. Although four options were given, no domain was ranked as 'not important'.

	Sample size	Mean	Std. Dev.	Minimum	Maximum
Physical Domain	82	3.77	0.425	3	4
Cognitive Domain	82	3.55	0.548	2	4
Social Domain	82	3.59	0.543	2	4
Affective Domain	82	3.40	0.606	2	4

Table1 Domain importance means and standard deviations

The chi squared test was run to check on any relationship between the number of teaching years and the importance given to the various domains.

			How long have you been teaching P.E.?		
			Less than 10 years	10 years or more	Total
Social Domain	Less important	Count	0	2	2
		Percentage	0.0%	5.1%	2.4%
	Important	Count	11	19	30
		Percentage	25.6%	48.7%	36.6%
	Very important	Count	32	18	50
		Percentage	74.4%	46.2%	61.0%
Total		Count	43	39	82
		Percentage	100.0%	100.0%	100.0%

Table 5 Experience vs Social Domain

$X^2(2) = 7.877$, $p = 0.019$

With a p value of 0.019 there seems to be a relationship between the number of years teaching and the importance given to the social domain with the younger cohort ranking the physical domain as very important (74.4%)

During the interviews, half of the teachers (n=4) also talked about the importance of social and cognitive outcomes that can be achieved in and through PE, alongside health and

physical outcomes. For two teachers, PE is '*more than*' just about the '*technical skills*' (T6, M). Rather, it was argued that PE is in a unique position to develop wider skills such as '*teamwork, respect, leadership, initiative, creativity'* mainly because PE teachers '*are not strictly bound by the syllabus as other subjects*' (T2, M). This gives teachers the freedom to '*create activities which can tap into those outcomes*' (T2, M). For another teacher, it was important to recognise that within the PE learning environment, there are pupils who are not good in sport but have developed social and cognitive skills. It was therefore important to support pupils to develop social and other skills:

> ***'It [Physical Education] is there to help the students learn social skills. Sometimes during a lesson you have good leaders even though they may not be that good in the sport we are doing. They may be good in organisation skills, or in leadership, helping each other out and teaching each other if they already have experience in the sport. For example if they already know basketball, they give tips to each other on where to stay on court. "I can't pass you the ball because there is someone else already".'***
>
> ***(T8, F)***

Six of the teachers interviewed identified four main domains that PE should aim to address, namely physical, social, cognitive and affective; and these four domains should also be assessed. However they remarked that doing so is much easier said than done, giving the impression that assessing all domains is more of an aspiration than a reality. But it was also apparent, as these teachers explained, that achieving this aspiration in practice was not always possible or feasible. The quote below illustrates some of the difficulties this teacher experienced when trying to assess these other domains, admitting that the physical domain is the most easily assessed.

'I think the physical domain is the easiest to assess. For example, if I have to assess whether the student knows how to do the high jump after the topic, it's quite easy to assess. On the other hand, if I had to assess, during the high jump topic, how much he is helping others, how much he is thinking about his mistakes and how to arrange them. Then it is a little more difficult to assess. But it is very important that they are.'
(T6, M)

Even so in the observations other domains were seen being tapped into.

From the observations there were three lessons that a clear emphasis on other domains than the physical were observed. In these lessons the students were split into small groups, with the teachers highlighting skills from other domains. Skills like social and affective domain where mostly tapped into. Following are some notes on two of the lessons observed:

Teacher: T5, M

Students: 13

Year: 7

Topic: Gymnastics

During gymnastics assessment students were split into groups of four doing skill analysis, an exercise on giving feedback, taking criticism and working as a team. Teacher asked for feedback on how it felt to be given feedback, being criticised and improving. An exercise in social and affective domains.

Teacher: T7, F

Students: 23

Year: 8

Topic: Badminton

During badminton assessment students were in groups of three, giving feedback to each other and assessing each other's capabilities. Even though at first glance the physical domain was seen to be assessed, the teacher was giving instructions on how to give feedback, receive it and help others achieve results. An exercise in the social domain.

Evidence from the teacher survey and interviews suggest that the content of teachers' assessment (i.e. what teachers assessed) was linked to some extent to what activity area was taught. However, overall, it was apparent that the physical domain was the area that was always assessed by the vast majority of teachers, as figures 5-8 suggest. For example, as shown in figure 5, as far as team games were concerned, more than 60% reported always assessing pupils' physical competence or performance. Similar results were obtained when individual activities and health-related fitness were taken into consideration.

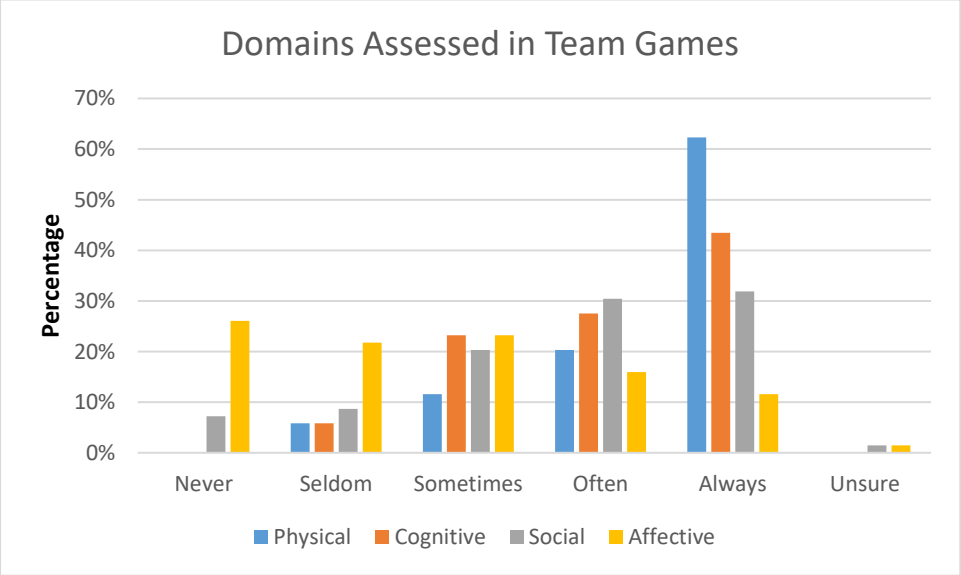

Figure 9 Domain ranking in the assessment of Team Games

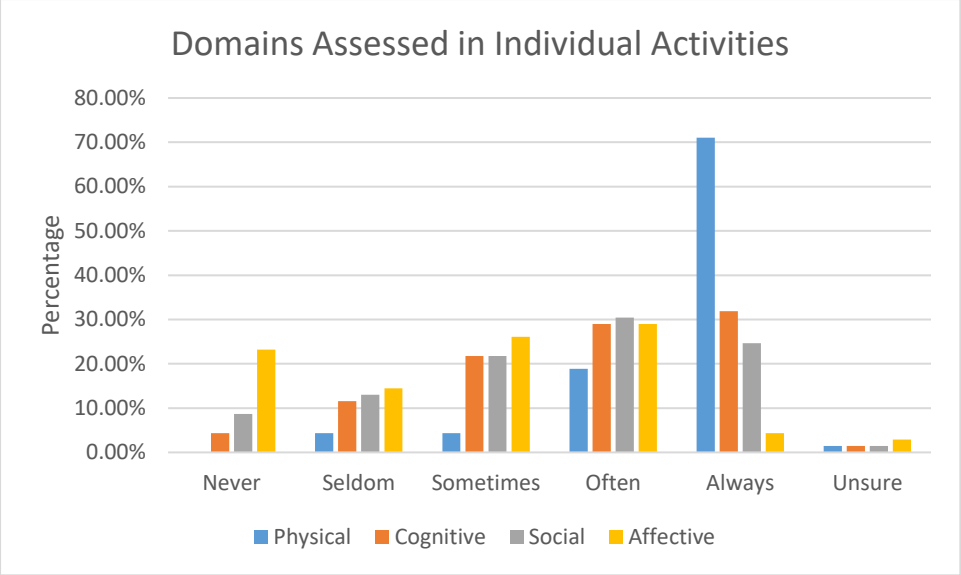

Figure 10 Domain ranking in the assessment of individual activities

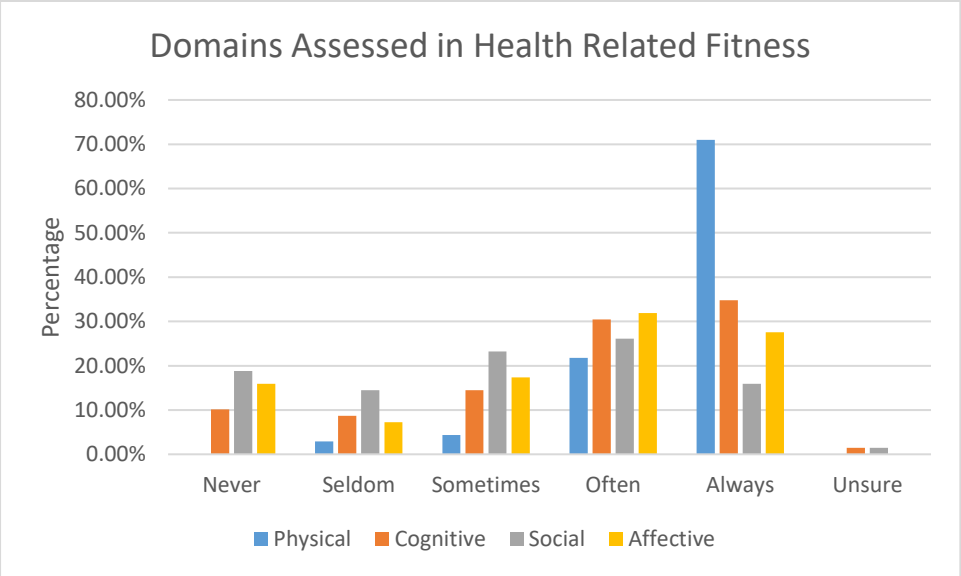

Figure 11 Domain ranking in the assessment of HRF

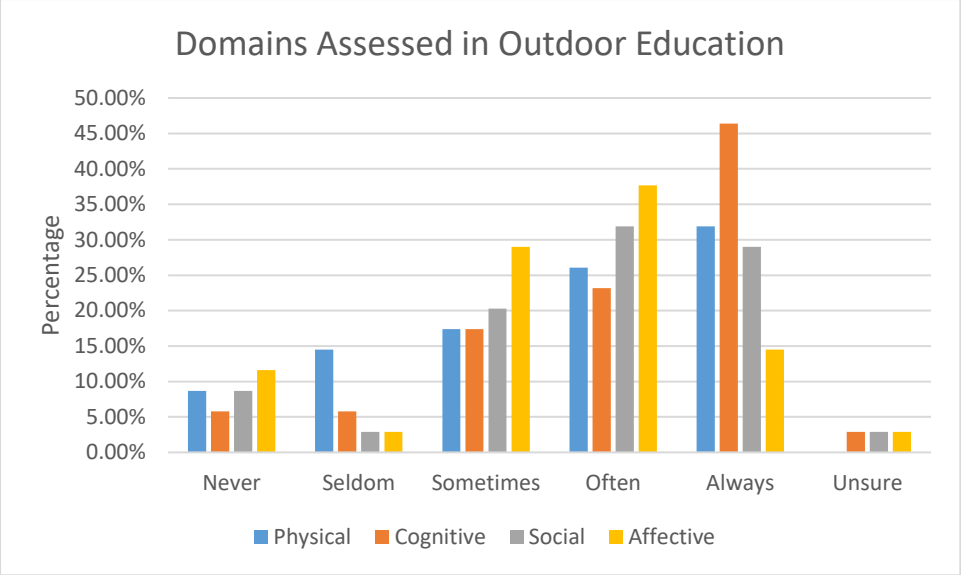

Figure 12 Domain Ranking in the assessment of Outdoor Education

One respondent made it very clear that it is very important that the students being observed know what they are being assessed upon.

> ***'It is done at the end of the 6 week course. I explain what the assessment criteria are. For example, you shall be playing a 4v4 game and I will be assessing the way you will be passing." The rest will be doing the same activity in set areas. I believe it is important that the students know what they are being assessed upon as they will focus on the task at hand and give more attention to the teaching cues regarding what is being assessed. I will then move from one group to another and observe each group on it. For example, shooting, jump shots, shooting. I sometimes takes 2 or more lessons.'***
>
> ***(T1, M)***

Some respondents highlighted the importance of assessing other areas than the physical domain, emphasising on the teacher's role as an educator.

> ***'I usually tell them that they have to appreciate that everyone has his own strengths and weaknesses. That you are capable of making it to a certain point, you are able to understand that your friend is better that you in certain things and less in other things, it is already a good step in social skills. More than that, if you can see that with your strengths you can help your friend, moreover how you can seek help to improve other skills. I think that would be the ideal development in the social domain. That student would gain the most marks.'***
> ***(T6, M)***

One respondent highlighted attitude and team work in particular.

> ***'I assess also attitude and other skills like bringing a team together. In PE, attitude is the most important thing. We as educators are here to look at the attitude, values, sharing. Even in interschool competitions, they should give trophies for encouragement, team work.'***
> ***(T5, M)***

In individual activities priority was given to assessing the physical domain. Emphasis was given to the fact that after a topic is over, assessment is done individually.

> ***'Many times, I assess as soon as I finish a topic. So, if I finish sprints or long jump i assess the following lesson. For sprints they have to run 60m. And I give marks according to the time that they manage to run in. I call them 2 at a time as I have 2 stopwatches and, on my signal, they sprint.'***
> ***(T2, M)***

Some respondents showed that they prefer assessing the skill as soon as it is done instead of waiting for the topic to end.

> ***'But when it comes to assess the children in the proper long jump it has to be individual. And that is very time consuming. I assess the distance that they jump. I have criteria based on the children's capacities and keep to it.'***
> ***(T7, F)***

> ***'In individual activities I assess every skill after it is done.'***
>
> ***(T1, M)***

One respondent emphasized that group work is still used even in the assessment of individual activities.

> ***'Then I divide them into four groups, they get the mats together, and they prepare a routine that involves all the skills that we covered. Then they are recorded by other students. All teams watch each other and we discuss discussing the videos, skills, creativity and all. While they are working I go around with my sheet, checking on teamwork, agility, etc. Sometimes they also do self-assessment. I give them sheets and they tick for each other's abilities.'***
> ***(T5, M)***

To avoid behaviour problems and off task behaviour some respondents adopt the strategy of giving a team game to the rest of the students, while they assess the rest.

> ***'I give them a ball while they wait so that they don't get bored and I don't have any behaviour issues.'***
> ***(T2, M)***

> ***'It is very difficult to assess a whole class in one way. I give them a football tournament and assess them [students] five at a time. You can give them individual attention focus on those five.'***
>
> ***(T1, M)***

The focus on the physical element was shown also in HRF. One respondent explained how the class was managed while doing the fitness test for a group of students.

> ***'As for assessment I do the sit and reach test, 6min run, etc. I split them in groups, one teacher takes the class for other activities and I test them.'***
>
> ***(T3, F)***

During the lesson observation students' interest to improve was clearly observed.

Field notes from observations

Teacher: T1, M

Students: 24

Year: 7

Topic: Health Related Fitness

[The teacher had the area purposefully set up to be able to conduct the fitness test. Student were split into groups of 6. The teacher went through the exercises to be performed. Already some students started to show interest. "Miss, how many laps did I manage last time?", "Miss what was my previous flexibility result? I have been stretching a lot at ballet!" They are feeling compelled to improve upon their own result. They are not motivated or interested to surpass each other, but to surpass their previous assessment result!]

Although the teachers interviewed lacked to assess outdoor education using phrases like 'We don't do outdoor here' (T7, F) through the survey one could see that the trend of focusing on the physical domain varied in Outdoor Education where the cognitive domain was given priority over the other domains (over 45%).

4.3 Barriers Encountered

Some teachers highlighted the problem of subjectivity in assessing Physical education which may result in unfair assessment for the learners:

> ***'Another thing is that assessment in PE can be very subjective. It depends on the examiner observer. It is very difficult to keep it objective. I might give a 7/10 while someone else might give a 5. A good assessment is one where whoever assesses comes up with the same or at least a similar mark. I think that I am more lenient than other teachers. It might be that it is because I don't believe so much in it.'***
> **(T1, M)**

One respondent also voiced his concern on reliability of assessment in team games had they to be done on an individual level.

> ***'I choose four criteria from attack and defence and I focus on them. I give a mark from 1 to 3 on that criteria. I will than give a mark on the frequency that I see a student do a particular item. Sometimes I doubt the fairness there as it is not possible to always observe the class. I sometimes consider assessing them individually, but I fear that the assessment will be tainted as it will not remain part of a game situation.'***
>
> ***(T2, M)***

There was a general perception that summative assessment was not conducted in a way to raise the status of PE in the eyes of this group of students:

> ***'I don't see that it has a place in PE, because those students who do not manage to arrive at the wanted point, they give up. Some do not even care about assessment. For some students having a mark in PE is not that important as they do not give the subject the same importance that they give to other subjects like English and maths.' (T3, F)***

In relation to this issue, one particular concern raised by four teachers during the interviews was on the fairness of the assessment, as marks may be given to students who are already have advantage in the particular topic being assessed.

> ***'One has to keep in mind that not all students are good in everything. If assessment is on one topic which may be the weakest link of the student it will not show the truth on the student. She can get a C in PE because assessment was on the topic she was weakest in. In fact I have started splitting assessment so that the parents know what has been covered and on what criteria the mark is based on.'***
> ***(T7, F)***

Some teachers found the emphasis on assessment knowledge against normative criteria and not learning progress highly problematic as captured in the quotes below:

> ***'…for example, in basketball dribbling I have a class where one child is the all-rounder, and is being assessed on dribbling. You give him a very high mark. You then have this other child who is obese, much less skilled. In the same term, Joe was skilful when you started and still skilful at the end. On the other hand the other child, started out knowing nothing about dribbling and at the end of the topic, he managed to dribble a ball around the court. I would love to give more marks to the last student to show him that he has done effort and tell to the other one to keep up the good work. For me I want to see progress. Not what you know, but what you've learnt. Unfortunately, we have a system which encourages knowledge and not learning.'***
> ***(T5, M)***

> ***'I do my best to assess fairly. If for example there is a student who does gymnastics afterschool, there is an unfair advantage on the other students who do not do. At the end of the day I am not a coach, I teach Physical Activity. Obviously that student will get full marks, but the rest will be assessed according to their competence. I think our assessment helps the students. They ask me, even in Parents' Day. They ask me, "75? How come that mark?" I then give them a breakdown of the mark. Many times I give feedback to the students after the lesson, or in the following lesson.'***
>
> ***(T8, F)***

One teacher showed particular concern regarding the time that the students have during assessment if there is a large discrepancy in possession of the ball in a game situation because of different skill levels.

'I try to be as fair as possible, but sometimes I feel that when I assess during a game I am not being that fair because you may observe a stronger student play as he has more possession while a weaker student may be observed less because he has less time with the ball and may rest on the other one to win the game.'

(T2, M)

For some teachers, not achieving well in summative assessments may have a detrimental effect on student motivation. This was explained by one teacher as follows:

'Students may find assessment in some cases as being unfair. This is mostly because they will get the mark of the term on a sport, say football, where some students have a big advantage because they train somewhere else. So during assessment those who train a different sport or do not go to a sport after school just get a low mark. They tell you "it's useless that you assess me, I won't do well'.

(T3, F)

Most teachers held a similar view on why assessment, if done well, was important. As the quote below illustrates, assessment is necessary to 'try and see what the students know and what they do not know' (T7, F) in relation to the objectives set by the teachers:

'First and foremost, assessment is done so that you see where the students stand and what they know. If I am assessing a particular topic, I have my own objective, and the assessment will show me whether the students have reached that objective or not, or where they are at the current moment in reaching the particular objective. For example, when assessing in sprinting I see how they compare the times on the assessment sheet.'
(T2, M)

When observing T2 during the assessment lesson it was very clear that he was comparing the results the students gained with the sheet given by the ministry. Feedback given to the students was on whether the desirable time was achieved.

Some respondents (n=3) immediately raised concerns at assessing a skill in a team game during a game-situation, admitting that they opt to do the team games assessment on individual skills. Time constraint was one of the reasons mentioned.

> ***'I take into consideration skills on their own. I have a sheet and mark the skills, and take notes.'***
> ***(T5, M)***

> ***'That [assessment in team games] is a problem as I would really like to assess the student during the game. You can do the best passes but get stuck in the game. Since in each class we are 28, it is very difficult. You must leave half the class standing if you do it in a game situation, while you do the assessment of the other class. I have to do it skill based and do different stations and they go around. It takes me 2 lessons to do this.'***
> ***(T7, F)***

One of the respondents mentioned a way that he tackles this by splitting the area into different stations.

> ***'I split them into stations where each group is trying out a particular skill, with one group doing the game. I go around the groups so that I can observe what each group is doing. Although I do leave a lot of space to the students, because if I had to assess each and every individual, I would end up not giving them the attention they require, as the time is quite limited.'***
> ***(T6, M)***

One respondent pointed out that assessing tactics can be hard because the students haven't mastered the individual skills yet.

> ***'I always tackled assessment in the form of an exam. I do it skill based. For example skills in volley ball, the set, the dig etc. It is much more difficult to assess during a game, or whether they have achieved good tactics, especially if they do not have enough sills to play the game well.'***
>
> ***(T3, F)***

Some respondents noted came forward as finding it hard to assess team games due to time and large number of students that have to be assessed.

'That [assessing team games] is a nightmare. I split a class of twenty into 5 groups and have small sided games. The problem with small sided games is for me to watch the students at all times. I will have the criteria in front of me. I then try to observe each student on the pre-set criteria.'
(T2, M)

Assessment can be quite tiring as it can take me up to 3 lessons at times. It may be that as I am assessing, I still don't manage to observe all the students. He might have done a good pass while I'm not looking. My assessment sheet becomes a nightmare as well. All with scribbles and markings. It takes a lot of time. (T2, M)

Time constraints seem to be a constant deterrent to assessment, with the participants feeling that authentic assessment could be at risk, making assessment unfair or unrealistic. In this case the PE teachers felt that assessing students, especially in team games can be a daunting task.

4.6 Summary

The aim of this chapter was to serve as a showcase of the main concerns, issues and ideas raised by the interviewees and to display the data collected in the surveys. Most teachers seem to be concerned by the current assessment practices, expressing concerns on the time it takes and on its effectiveness. A big emphasis was made on the importance of the type and quality of feedback given. Both phases of data collection illustrated other valid information as pedagogies and teacher experiences. The next chapter is intended to tackle the key points arising in the analysis chapter by drawing on international and national literature to give more insight to the main issues and concerts that surfaced.

Chapter 5. Discussion

5.1 Introduction

In chapter five the results from all three modes of data collection were gathered and presented in accordance to their common themes. This chapter seeks to look across these themes and discuss what the results mean and how they relate to the local policy contexts and the wider literature on assessment in education and physical education. This chapter is split in accordance to the research questions and will also highlight and suggest any limitations and directions for further research.

5.2 What are PE teachers' understandings of the notion of assessment and assessment policies in Malta?

Some respondents highlighted that assessment, as currently practiced in schools was primarily about allocating a grade to students; that is, capturing where the learners stand at a particular point in time and what knowledge they display. Summative assessment is widely acknowledged as the process of assigning a grade at the end of a 'scheme of work'. This grade provides evidence of student progress and the teachers acknowledged that this was often used in schools as a 'proof' about the extent to which the students have achieved the criteria or not. They explained the emphasis placed on communicating this grade to parents (without always detailed explanations however) as part of their increased accountability.

The fact that most of the teachers who participated in the interviews appeared to understand assessment as being primarily a summative process, reflects the importance contemporary educational discourse in Malta is placing on measuring whether the learning outcomes have been achieved (Stobart, 2008). This was also combined with the common perception that assessment was promoted as part of the bureaucratic protocol of the schools.

Some of the teachers interviewed showed a degree commitment in that process; some believed that it was important to confirm what students knew and were capable of doing in relation to curriculum outcomes at the point of assessment (Gibbons and Kankkonen, 2011; Harlen W. , 2004). In this context, it is important to remember Dylan William's reflections that emphasise the importance of understanding where students are in their learning:

> "A mismatch between a child's capabilities and what he or she is being asked to do is disastrous. When the level of competence is high, and the level of challenge is low, you get boredom, and when the level of competence is low, and the level of challenge is high, you get alienation. But when the level of challenge is just at the limit of your competence, you get this feeling of flow" (Williams, 2006)

Yet, what William's emphasises is the importance of this type of assessment that happens all the time to inform teachers' decision making. In the context of the present study, some teachers had concerns about the summative assessment processes as currently experienced as lacking real meaning and, quite problematically, as having a potentially negative effect on students' confidence. The fact that not being able to achieve well in summative assessments may have a negative effect on student motivation is a known fact (Rodriguez, 2004) with effects on achievement, attitude and motivation too. Some teachers also feared that an over-reliance on summative assessment might hinder the teaching and learning process. In line with Zigo and Moore's (2002) argument, some explained that implementing tests and other forms of summative assessments required substantial time and restricted the available time for teaching and learning.

.

Most teachers argued however that, if done well – in a valid and reliable way, summative assessment does actually give a realistic view of the students' progress. Less than half of the interviewed teacher believed that beyond its narrow use, assessment should be used as part of teaching and learning. These teachers' argument were aligned with contemporary understandings of the notion and importance of Assessment for Learning (AfL). Formative assessment or AfL, as explained in the literature review, is frequently referred to as 'a continuous process rather than a discrete event' (Frapwell, 2010, p. 108) that supports pupils to understand where they are in their learning, where they need to go next and how best to get there.

The Maltese National Curriculum Framework (NCF) clearly states that assessment should be used as a valuable tool to enhance learning (Ministry of Education and Employment, 2012). It is important to note that discussions about the use and importance of AfL arose in the context of teachers' dissatisfaction with an overlying on summative assessments. They believed that having a summative mark at the end of a topic was not being enough for the student to improve and succeed. Respondents in the interview mentioned the use of AFL as a tool to make sure learning is taking place in an active manner. AfL can be considered as being an active process (Hay and Penney, 2009). This was expressed by various participants and compares to what Hay and Penney (2009) conclude that information collected in the assessment process should help structure future learning.

More recently, Tolgfors (2018) suggested that AFL can be an effective way to ignite 'physical activation' (p.137) where the students are empowered to take responsibility for their own learning. One respondent highlighted the way AFL helped the students in starting talking and discussing their learning and understanding. The importance of providing good

constructive feedback for improvement was also highlighted by many teachers in the present study. The use of feedback is acknowledged as one key AfL strategy in the international literature. It has been explained that the main purpose of feedback is to 'reduce discrepancies between current understandings and performance and a goal' (Hattie and Timperley, 2007, p. 3). In this context, as Hattie and Timperley (2007) explain, the main goal is to fill the gap between what students currently do and understand and what they need to do and understand to progress in their learning.

It is interesting to note most teachers seemed to interpret feedback as information regarding of students' performance or understanding provided by them, the teachers. Only a small number of teachers, as discussed later in this section, acknowledged that feedback can be provided by other agents, such as peers.

Some of the key messages conveyed by some of these teachers were in line with developments in the role and importance of assessment internationally. Carroll (2005), for example, amongst others (Black and William, 1998; Hay and Penney, 2009; Black, McCormick, James, and Pedder, 2006), advocates that assessment is an essential part of the process of teaching and learning. In the context of PE, Mosston and Ashworth (2002) describe assessment as being universal and present in all aspects of life. The fact that half of the teachers acknowledged the importance of assessment as embedded in teaching and learning also reflect a significant departure from findings reported in 2016.

In 2016, a local PE study (Falzon, 2016) sought to identify factors that affect the quality of physical education provision. One factor explored was the frequency of use of assessment by PE teachers. In the context of this study, and based on the perceptions of PE teachers involved, assessment was rarely perceived to be a continuous. In fact, the preferred method

of assessment at the time of the study was the end of term or end of unit assessment. In the context of the present study, the majority of teachers have indeed stated that they used mostly end of the topic assessments.

Another important set of results reported in the previous section relate to what PE teachers should assess. The multi-dimensional nature of Physical education was acknowledged by most teachers interviewed, with teachers identifying not only the importance of support pupils' physical competence but also their cognitive development (i.e. understanding), affective domain and social development. Some argued that these were important aims to achieve within PE lessons.

5.3 What assessment practices / approaches do PE teachers employ and why?

It was however also evident that most interviewees suggested that the physical domain was prioritised both during the lessons and for assessment purposes. This was supported by the results from the survey. A breakdown of assessment practices per activity area or subject suggests that in most cases, physical competency was prioritised. Some teachers explained that despite their commitment in supporting learners develop cognitively, emotionally and socially, assessing these dimensions was sometimes an impossible task, as they are not as easily assessed as the physical domain.

In some ways, this is not a surprising finding. Academics for decades have argued that assessment practices in PE prioritise the execution of particular skills (Hay and Penney, 2009). It appears that this is the case of assessment in PE in Malta too. Interestingly, the assessment sheets provided by the Ministry of Education seem to encourage observation and marking of physical and cognitive skills attribute over social and affective ones. Laker's (2001) advocacy on the importance of teaching the different domains, especially in teaching skills like

creativity, commitment and team work is still relevant today. Yet, the suggestion that it is important to not only seeking to develop these traits, but to assess them seems to be a much more daunting task for the PE profession.

The teachers who participated in the present study appeared to draw upon some common strategy as far as assessment was concerned. The way they assessed also appeared to reflect the way they taught an activity. The case of games provides a useful illustration. As far as team games were concerned, the interviewees seemed split between command style of teaching and Teaching Games for Understanding (TGFU) with one teacher quoting guided discovery as the mode of teaching this topic. It was clear that there was a direct link between the method of teaching and the approach used for assessment. Those reporting using TGFU also mentioned assessing pupils while working in small groups. Those using a more traditional approach (the command style) went for assessing individual skill development as students were performing those skills individually, one student at a time.

In relation to what was assessed, results from the survey also suggest that, in the case of teaching team games, the physical domain was given the most importance during assessment followed by the cognitive domain. This was also the case for Health related fitness and Individual activities. This shows that the physical domain remains the most domain to be given importance throughout the topics.

Whether against a set criterion or just to provide students feedback, the majority of teachers in both the survey and the interview stage mentioned that observation was the primary means used to collect evidence on student progress. This was seen across all areas in physical education. The assessment sheets used in states schools and some church schools seem to lead to this type of assessment, as a tick box format assessment sheet is provided to

the teachers with a lot of emphasis on the physical domain. Laker (2011) mentions that this [observation] type of assessment is authentic adding that this adds to the quality of assessment both for evaluating the effectiveness of teaching and for the detailed progress of the individual student.

It is interesting to note that two respondents emphasised the use of technology in the assessment process. Through video analysis and assessment for learning the respondents highlighted how technology can assist both teacher and student in the process of learning and assessing. They emphasised that technology should be part of teachers' practice as a tool to enhance the teaching and learning experience. This is also acknowledged in the relevant literature. It has been argued that technology should be used to enhance what teachers do and to improve the quality and impact of the educational process (Juniu, 2011; Casey, Goodyear, and Armour, 2016). As the two teachers claimed, the use of technology such as tablets to enable students to observe their performance with the aim to improve it was paramount. Eberline and Richards (2013) have also mentioned the used of filming the students so that they can examine and assess what they are doing, thus giving the teachers the support they need to enhance the learning process.

During the observation sessions some teachers showed that the use of traditional methods of assessment were still preferred to technological devices. Although this study doesn't delve into this issue, Kretschmann's (2015) study suggested that over 80% of the participants did not have enough knowledge on how to integrate technology into the teaching process, with the majority of teachers preferring more traditional technologies such as blackbirds and images. Casey et.al (2016) emphasised on the need for educators in Health and Physical Education to enthusiastically use technology in our practices as it has become

part and parcel of the cohort of students that are faced in schools. This is thus potentially another important new direction for PE teachers in Malta, especially when further research evidence confirms the importance and impact of such approaches to student learning.

Self- and peer assessment ranked as the least favourite assessment strategy in the survey. During the interview stage, only a few respondents mentioned the application of peer assessment procedure as a good approach that can enhance learning. The use of self- and peer-assessment approaches is not a new concept. Almost twenty years ago, Jonson (2001) acknowledged that when peer assessment is used for formative assessment it can boost the overall achievement and experience of learning. Since then, there is accumulative evidence on the benefits of these AfL strategies (ibid.).

As well as developing motor skills, using peer assessment can also assist in developing critiquing skills, criticism, tolerance and acceptance (Johnson, 2001; Johnson, 2004; Gibbons and Kankkonen, 2011). Moreover, it can provide students with immediate feedback on the task at hand. Yet, as previously noted, the use of this AfL strategy did not appear to be widespread amongst the PE teachers involved in the present study. One PE teacher interviewed appeared to be an advocate of self- and peer assessment. He however also acknowledged that students are not always adequately prepared to engage in that process effectively. In that case, the teacher felt he had to 'intervene' to ensure that the process of giving and receiving feedback was done effectively.

5.4 What barriers do PE teachers encounter that hinder their assessment practices?

Overall, the results suggest that these teachers needed further training in developing their understanding of the various ways they can use assessment to maximise student learning. Evidence from this study suggest that no specific training on assessment was given at university level. This could explain why a limited range of AfL strategies were employed by these teachers. As Popham (2009) suggests, a lack of, or insufficient know how on assessment and its practices may diminish the quality of education. He however continues that developing competency on and about assessment is essential for the modern educator.

In the context of PE research, this lack of training was noted by Veal (1988), *who* argued that one of the reasons that restricted PE teachers from engaging in formal assessment in meaningful ways was the limited support they had during initial teacher education. This still seems to be the case for these Maltese PE teachers. More recently, Taras (2007) and Shepard et al. (2005) have also argued that, overall, assessment was a neglected area of study in initial teacher education programme. This is supported by DeLuca and Klinger (2010) who noted that new teachers in particular were quite unprepared for assessment practices. This seems thus to be a recurring problem across time and national contexts.

There is also of course the issue of the ongoing, career-long support teachers have about assessment. DeLuca and Klinger (2010) observed that educators' skills and knowledge on assessment is quite low throughout their career. Once appointed, although teachers have opportunities to share ideas and practices with colleagues (Johnson, 2013), a lack of formal, structured support, as appeared to be the case for these Maltese PE teachers, is problematic.

With the implementation of the new leaning outcomes and lack of guidance throughout their careers, teachers may find it difficult to improve their assessment performance.

Beyond this limited access to high quality initial teacher education and Continuing Professional Development (CPD), these PE teachers identified some practical barriers in their endeavours to engage in assessment. For example, some argued that there wasn't enough time to conduct assessment in a desirable manner. Teachers also mentioned that having a large number of students in class didn't help the cause. Both Torrance and Pryor (2001) and Collier (2011) explained that practical barriers such as these identified by these teachers are often the main reasons that inhibit new and various forms of assessment practices to be implemented. This is a significant consideration that should be taken into account in efforts to educate teachers about assessment. Other such practical barriers include available facilities, as identified by Kinnunen and Lewis (2013). But there is also the barrier in changing teachers' understanding – and even perceptions - on the importance of assessment. In the case of the present study, some teachers expressed concern that assessment is somewhat of a burden. It could be the case that assessment expectation in Malta are justified in the context of increasing accountability. To lessen the burden teachers currently feel, it is important to enhance teachers' understanding on the importance of assessment and to emphasise assessment as a tool for learning rather than a tool for accountability purposes.

5.5 Limitations of the Study and Directions to Future Research

This study is not of course without its limitations and these need to be discussed. Phase one relied primarily on qualitative methodologies. Qualitative research has been criticised for being 'subjective'. It is important to underline that both phases of the research had a subjective element as both the interview questions and observation protocol as well the

questions set in the survey questionnaire (despite dominated by 'closed', quantitative questions) were influenced by the student researcher's existing knowledge, perceptions and engagement with the relevant literature. So, it is important to underline that the student researcher does not seek to claim absolute objectivity and independence. However, it is also important to state that a number of strategies were employed to ensure that the study was rigorous and credible. For example, pilot studies were conducted for all three data collection tools.

Another limitation arose in the course of the study as due to limited resources and lack of time, it was not possible to collect the reflections of the participants. To make up for this the researcher engaged with the participants as much as possible, asking them to clarify any issues that were unclear or probing with their own response. In this manner the researcher ensured that the interpretations were as accurate as possible. To try and limit the issue of subjectivity the respondents were later handed a copy of the results. One common limitation pertaining to the three instruments is 'reactive' or 'observer' effects (Denscombe, 2003, p. 53; Johnson and Turner, 2003, p. 304) as there was always the fear that teachers could be inclined to present a distorted image of reality and report only what they deemed as being socially desirable thus altering their behaviours or answers.

Phase one was also small scale. In the case of small scale, case study research, lack of generalisability in the traditional sense is perceived to be problematic. To counteract this, researchers have proposed the importance of maximising the possibility of case-to-case transferability. This means that by providing details, rich contextual information, readers can learn from the case. However, in retrospect, this was not done to a sufficient level of detail.

Future research should thus examine the relevant school policies, talk (potentially informally) to other PE teachers (and teachers from other subjects), attend meetings etc., with a view to develop a deeper understanding of the broader contextual influences that might shape teachers' perceptions. Such information would also allow the student researcher to explain why certain perceptions were held and why the observed practices were evident.

To address the limitation of the lack of generalisability, a decision was made to develop and administer a national survey to a 'representative' sample of PE teachers. The final response rate was good helping to give weight to the results found. Another limitation was that no interviews were held after the survey. This would have had ironed out any further questions arising from the survey. Finally, future research should also examine students and their own perceptions of the assessment practices they experience as part of their PE experiences. Other areas for future research would be to explore the ways technology could help in the process of assessment for learning, and also what assessment practices are being held in other year groups and the initiation of the Learning outcomes progresses through the years.

Chapter 6. Conclusion

The present study set out to explore the nature, range and efficacy of physical education assessment practices in secondary Maltese schools. Specifically, the study sought to answer the following research questions; (i) What are PE teachers' understandings of the notion of assessment and assessment policies in Malta?, (ii) What assessment practices / approaches do PE teachers employ and why?, and (iii) What barriers do PE teachers encounter that hinder their assessment practices? The overall aim was to collect robust evidence in order to inform both policy and practice.

Using mixed methods, the study was conducted in two phases. The first phase focused on examining PE teachers' perceptions and practices about assessment. A total of eight PE teachers agreed to participate in the first phase the study. The primary data collection tool was semi-structured interviews supplemented with lesson observations for the purposes of triangulation. Results from phase one informed the development of a national teacher survey (phase two). This was designed to explore teachers' perceptions and understanding of the purpose and application of assessment on a larger scale. Important questions around teachers' experiences and views on their initial teacher education and professional development opportunities in relation to assessment were also posed. The survey was distributed to all secondary PE teachers employed in state and church schools nationwide, and returned by 90, which represents a 71% response rate.

Results reported in this thesis suggest that there are a number of important areas which need to be addressed. Firstly, it is important that PE teachers are aware of and acknowledge the existence of a wide range of assessment processes and procedures, including summative and formative (or assessment for learning) assessments. It is particularly

important that they share a more in-depth understanding of the importance and complexity of AfL. To address this, emphasis should be placed on the content and quality of initial teacher education and CPD as it appears that they currently neglect this important aspect of teaching and learning.

Implications and Recommendations for practice

PE teachers are the ones responsible to facilitate learning and implement assessment in the subject. It is clear that the notion of formative assessment still needs time and training to be fully understood, conceptualized and delivered. It is very clear that assessment and its practices need to be part of both the initial and continuous teacher training. It is vital that current study units revolve around, or at least give weight to assessment practices in each specific area.

Although PE teachers tend to collaborated as individual departments in schools, sharing of ideas and resources for assessment practices would help teachers not to feel isolated when facing barriers that inhibit assessment practices. It would also provide enough grounds to provide teacher training and discussion groups to help tackle any of the existing barriers. With the implementation of the new learning outcomes, it is very important that the PE teachers are given tools to deal with assessment in a formative and holistic manner.

Appendices

1.Interview Protocol

Interview Protocol

Interviewee: _____

Interviewer: Karl Cortis

Topics

Discussed:_____

_____

_____

_____

Assessment Types and Formats interviews

Introductory Paragraph

To facilitate note-taking, I would like to audio record our conversation today. Please sign the release form. For your information, only my tutor and I will have access to the recordings which will be eventually destroyed after they are transcribed. In addition, you must sign a consent form. Essentially, this form states that: (1) all information will be held confidential, and (2) that your participation is voluntary and you may stop at any time if you feel uncomfortable. Thank you for your agreeing to participate.

This interview should last no longer than one hour.

Introduction

You have been selected today because you have been identified as someone who has a great deal to share about assessment. The research project as a whole focuses on the shedding light on assessment practices in this country and its effectiveness on student learning. The study

does not aim to evaluate your techniques or experiences. Rather, it is meant to shed light on what is happening and how good practice is being conducted.

Part A:

Theme: Interviewee background

How long have you been teaching Physical Education?

Which year groups did you experience teaching?

And for how long?

Which topic do you find most difficult to teach in PE? Why?

Which topic do you prefer teaching? Why?

Part B:

Theme: Teacher's views on assessment

What do you understand by the term assessment?

What is its place in Physical Education?

What should be assessed in Physical Education?

What do you understand by the terms Assessment of learning and Assessment for learning?

Do you make use of them in Physical Education?

Part C:

Theme: Assessment training and professional Development

What type of assessment training did you receive at university?

Did you ever attend courses regarding assessment in PE or otherwise?

Did you ever discuss assessment as a team?

What are the departments view on Assessment in Physical Education?

Part D:

Theme: Teaching styles in relation to assessment practices

There are four groups, (i) Games, (ii) Individual Activities (such as Badminton and Athletics), (iii) Health-Related Fitness, and (iv) one other optional activity (taken either from Games, Individual Activities or Outdoor Education) in PE.

What teaching style would you describe as using when teaching Games? Example? (Command, reciprocal, exploratory, etc.)

When it comes to its assessment, how do you assess team games?

What teaching approach you describe as using when teaching Individual Activities? Example? (TPSR, TGFU, Skill-based, etc.)

When it comes to its assessment, how do you assess individual games?

What teaching approach you describe as using when teaching Outdoor Education? Example? (TPSR, TGFU, Skill-based, etc.)

When it comes to its assessment, how do you assess Outdoor Education?

Part E

Theme: Effectiveness of assessment

How do students react to assessment?

Does assessment effect student's learning?

2.Observation Schedule

School:	Date:	Time:
Year Group:	Venue:	Number of Students:
Teacher Code:		

Topic: _____

Learning Outcomes: _____

General Description of Assessment lesson:

Main assessment strategy: _____

Time teacher spent explaining assessment procedure: _____

Time on task: _____

On task student communication: _____

Teacher feedback: _____

Student feedback: _____

Time spent by teacher doing assessment per student: _____

3.Survey

Assessment in Physical Education

As part of my research I am investigating assessment practices in Physical Education in Maltese Secondary Schools. To this end I have designed a survey to:

- elicit the ways assessment is done in the secondary setting;
- explore any correlation between assessment practices and pedagogies used; and
- explore any correlation between assessment practices and topics being assessed.

Please fill in all the questions presented in this questionnaire. All procedures have been approved by the University of Birmingham Ethical Review Committee and all information provided will be treated in strict confidence. The questionnaire is anonymous, thus your confidentiality is assured.

This study is being carried out by Karl Cortis under the supervision of Dr. Kyriaki Makopoulou from the University of Birmingham. This study is part in fulfilment of an MSc by research in Physical Education and Sport Pedagogy.

There are 36 questions in this survey

Part 1 - General information

[]What is your age? *

Please write your answer here:

[]Gender *

Please choose all that apply:

- ☐ Male
- ☐ Female
- ☐ Other

[]How long have you been teaching Physical Education? *

Please write your answer here:

[]Which type of school do you teach in? *

Please select at most one answer

Please choose all that apply:

- ☐ State School
- ☐ Church School
- ☐ Independent School

[]Please indicate the highest qualification that you have obtained. *

Please select at most one answer

Please choose **all** that apply:

- ☐ PHD
- ☐ Masters Degree
- ☐ Teaching Qualification MTL
- ☐ Bachelor in Education (Physical Education)
- ☐ Bachelor in Education (Other Subject)
- ☐ B.Sc in Sport
- ☐ Teaching Quallification PGCE
- ☐ MCAST Sports Diploma
- ☐ Other

[]What year groups are you currently teaching? *

Please choose **all** that apply:

- ☐ Year 7 (Form 1)
- ☐ Year 8 (Form 2)
- ☐ Year 9 (Form 3)
- ☐ Year 10 (Form 4)
- ☐ Year 11 (Form 5)
- ☐ Other: ____

[]In what way would you rank the domain importance in Physical Education? *

Please choose the appropriate response for each item:

	Very Important	Important	Less Important	Not important at all
Physical Domain	○	○	○	○
Cognitive Domain	○	○	○	○
Social Domain	○	○	○	○
Affective Domain	○	○	○	○

[]What is your current teaching status? *

Please choose **all** that apply:

- ☐ Regular teacher
- ☐ Part time teacher
- ☐ Supply

Part 2 - Assessment Training

This section is about any training and preparation you have had regarding assessment.

[]Did you receive any training at universtiy regarding assessment in general PE and how to conduct it? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

[]If yes, what type of training did you receive? (title, duration, etc)

Please write your answer here:

[]Did you ever attend any courses regarding assessment in General PE? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

[]

If yes, what type of training were you given? (training type, duration, provider, location, year)

Please write your answer here:

[]How effective would you rate these courses?

Please choose the appropriate response for each item:

	Very Effective	Effective	Barely Effective	Not Effective
Training at university	○	○	○	○
Other Training	○	○	○	○

Part 3 - Assessment Type

Assesment Type

[]Which Syllabus do you follow? *

Please choose **all** that apply:

- ☐ National Syllabus
- ☐ Own School Syllabus, based on the National Syllabus
- ☐ Own School Syllabus, set up comletely different from the National Syllabus

[]Which assessment scheme do you use for PE general? *

Please choose **all** that apply:

- ☐ None
- ☐ National set assessment
- ☐ School set assessment (assessment sheets are created by the PE department of school)
- ☐ Other: _____

[]When does assessment usually take place? *

Please choose the appropriate response for each item:

	Always (In every activity)	Frequently (In the majority of activities but not all)	Rarely (In one or more activities but less than half)	Never (In no activities at all)
At the start of the topic	○	○	○	○
During the topic (ongoing)	○	○	○	○
In the middle of a topic	○	○	○	○
At the end of a topic	○	○	○	○

[]What do you assess in Team Games? *

Please choose the appropriate response for each item:

	Never	Seldom (In some of the times I assess but less than half)	Sometimes (In approximately half the time I assess)	Often (The majority of times I assess but not all)	Always (Every time I assess)	Not Sure
Physical Domain (Skills, Physical Literacy, Techniques)	○	○	○	○	○	○
Cognitive Domain (Game Play , Tactics)	○	○	○	○	○	○
Social Domain (Team work, participation effort)	○	○	○	○	○	○
Affective Domain (Feelings and Emotions)	○	○	○	○	○	○

The Affective domain refers to student's feelings, attitudes and values. Students' value for movement, morality and ability to express their feelings.

[]What do you assess in Individual Activities? *

Please choose the appropriate response for each item:

	Never	Seldom (In some of the times I assess but less than half)	Sometimes (In approximately half the time I assess)	Often (The majority of times I assess but not all)	Always (Every time I assess)	Not sure
Physical Domain (Skills)	○	○	○	○	○	○
Cognitive Domain (Game play, tactics)	○	○	○	○	○	○
Social Domain (Effort, Participation)	○	○	○	○	○	○
Affective Domain (Feelings and Emotions)	○	○	○	○	○	○

[]What do you assess in Health Related Fitness? *

Please choose the appropriate response for each item:

	Never	Seldom (In some of the times I assess but less than half)	Sometimes (In approximately half the times I assess)	Often (The majority of times I assess but not all)	Always (Every time I assess)	Not Sure
Physical Domain(Stamina, Strength, Speed, Suppleness)	○	○	○	○	○	○
Cognitive Domain (Understanding general concepts of fitness)	○	○	○	○	○	○
Social Domain	○	○	○	○	○	○
Affective Domain (Improving own results)	○	○	○	○	○	○

[]What do you assess in Outdoor Education?

Please choose the appropriate response for each item:

	Never	Seldom (In some of the times I assess but less than half)	Sometimes (In approxiamtely half the time I assess)	Often (The majority of times I assess but not all)	Always (Every time I assess)	Not Sure
Physical Domain (Skills)	○	○	○	○	○	○
Social Domain (Team work, leadership)	○	○	○	○	○	○
Cognitive Domain (Critical Thinking)	○	○	○	○	○	○
Affective Domain (Feelings and Emotions)	○	○	○	○	○	○

You may opt not to reply to this question if Outdoor Education is not done in school, or is not assessed.

[]What happens to the assessment results? *

Please choose the appropriate response for each item:

	Never	Seldom	Sometimes	Often	Always
Shared with the students	○	○	○	○	○
Shared with School Administration	○	○	○	○	○
Shared withing the PE Department	○	○	○	○	○
Shared with Parents	○	○	○	○	○

[]Which type of assessment methods do you use in Team Games? *

Please choose the appropriate response for each item:

	Never	Seldom	Sometimes	Often	Always
Observation and feedback	○	○	○	○	○
Observation and comparing against a set standard of criteria	○	○	○	○	○
Observation and giving a mark on what is observed	○	○	○	○	○
Student Self - Analysis	○	○	○	○	○
Student Partner - Analysis	○	○	○	○	○

[]Which type of assessment methods do you use in Individual Games? *

Please choose the appropriate response for each item:

	Never	Seldom	Sometimes	Often	Always
Observation and feedback	○	○	○	○	○
Observation and comparing against a set standard of criteria	○	○	○	○	○
Observation and giving a mark on what is observed	○	○	○	○	○
Student Self - Analysis	○	○	○	○	○
Student Partner - Analysis	○	○	○	○	○

[]Which type of assessment methods do you use in Health Related Fitness? *

Please choose the appropriate response for each item:

	Never	Seldom	Sometimes	Often	Always
Observation and feedback	○	○	○	○	○
Observation and comparing against a set standard of criteria	○	○	○	○	○
Observation and giving a mark on what is observed	○	○	○	○	○
Student Self - Analysis	○	○	○	○	○
Student Partner - Analysis	○	○	○	○	○

[]Which type of assessment methods do you use in Outdoor Education?

Please choose the appropriate response for each item:

	Never	Seldom	Sometimes	Often	Always
Observation and feedback	○	○	○	○	○
Observation and comparing against a set standard of criteria	○	○	○	○	○
Observation and giving a mark on what is observed	○	○	○	○	○
Student Self - Analysis	○	○	○	○	○
Student Partner - Analysis	○	○	○	○	○

Part 4 - Student Learning

[]In your opinion, does assessment-as currently practiced in your shcool- have a positive effect on students' learining? *

Please choose **all** that apply:

- ☐ Yes
- ☐ No
- ☐ Sometimes

[]If you answered YES or SOMETIMES in the previous question, please explain why and provide an example of positive impact on students learning.

Please write your answer here:

[]

If you answered NO in the previous question, please explain why.

Please write your answer here:

[]

To what extent, does information collected through the assessment process- as currently practiced- is used to support students' progress?

*

Please choose all that apply:

- ☐ Not at all
- ☐ Slightly
- ☐ Moderately
- ☐ Quite a bit
- ☐ A great deal

[]To what extent, does assessment give a real picture of the students' progress?
*

Please choose all that apply:

- ☐ Not at all
- ☐ Slightly
- ☐ Moderately
- ☐ Quite a bit
- ☐ A great deal

Progress refers to the students' advancement from his/her initial point and getting closer to the target objective.

[]If you answered moderately or more, please explain why.

Please write your answer here:

[]To what extent, does assessment give a real picture of what the student has learnt? *

Please choose **all** that apply:

- [ ] Not at all
- [ ] Slightly
- [ ] Moderately
- [ ] Quite a bit
- [ ] A great deal

This refers to whether a student has reached the final objective or not.

[]If your answered moderately or more, 4 please explain why.

Please write your answer here:

[]To what extent, do you think current assessment practicest effect students motivation? *

Please choose **all** that apply:

- [ ] Not at all
- [ ] Slightly
- [ ] Moderately
- [ ] Quite a bit
- [ ] A great deal

[]In what way do current assesment practices effect the students' motivation? *

Please choose all that apply:

- [ ] Yes in a positive way
- [ ] Yes in a negative way
- [ ] Sometimes in a positive way
- [ ] Sometimes in a negative way
- [ ] No not at all

[]If you answered Yes or Sometimes in the previous question, please explain why.

Please write your answer here:

References

Allred, S. B., & Ross-Davis, A. (2011). The Drop-off and Pick-up Method: An Approach to Reduce Nonresponse Bias in Natural Resource Surveys. *Small-scale Forestry, 10*(3), 305-318.

Andrade, H., & Valtcheva, A. (2009). Promoting Learining and Achievment Through Self-Assessment. *Theory into Practice Vol.38, no.1*, 12-19.

Angelo, T. (1995). Reassessing (and defining) assessment . *AAHE Bulletin*, 7.

Apps for Physical Education. (2018, March 22). Retrieved from Support real teachers: https://www.supportrealteachers.org/iphoneipad-apps.html

Assessment, C. f. (2007). *Ashfield Girl's School: Peer and Self Assessment in Physical Education.* Belfast: CEA.

Avery, M. (2012). Web-based assessment of physical education standards. *Jounal of Physical Education, Recreation and Dance, 83*(5), 27-34.

Beatty, P. C., & Willis, G. B. (2007). Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quaterly, 71*(2), 287-311.

Bennet, S., Maton, K. A., & Kervin, L. (2008). The 'digital natives' debate: a critical review of the evidence. *British Journal of Educational Technology, 39*(5), 775-786.

Bennett, R. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy, and Practice, 18*(1), 5-25.

Berstein, B. (1971). *Class, Codes and Control. Volume : Theoretical Studies Towards a Sociology of Language.* London and Boston: Routledge and Kegan Paul.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into Practice.* Berkshire: McGraw-Hill.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: a theoretical inquiry. *Research Papers in Education, 21*, 119-132.

Brooksie, C. (2014, August 25). *17 ways to use augmented reality in #physed*. Retrieved from Physedagogy: https://physedagogy.com/2014/08/25/17-ways-to-use-augmented-reality-in-physed/

Butt, G. (2010). *Making Assessment Matter.* London: Continuum International Pub. Group.

Carroll, B. (2005). *Assessment in Physical Education.* London: The Falmer Press.

Casey, A., Goodyear, V. A., & Armour, K. M. (2016). Rethinking the relationship between pedagogy,technology and learning in health and physical education. *Sport, Education and Society, 22*(2), 288-304.

CERI. (2008). *Learning in the 21st Century: Research, Innovation and Policy.* Paris: OECD.

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: the use of a midpoint on the likert scale. *Performance improvement, 56*(10), 15-23.

Claxton, G. (1995). What kind of learning does self-assessment drive? Developing a 'nose' for quality: comments on Klenowski (1995). *Assessment in Education, 2*(3), 339-343.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education.* New York: Routledge.

Coleman, M. (2012). Interviews. In A. R. Briggs, M. Coleman, & M. Morrison, *Research Methods in Educational Leadership & Management* (pp. 250-265). London: Sage Publications Ltd.

Collier, D. (2011). Increasing the value of physical education. *Journal of Physical Education, Recreation and Dance, 82*(7), 38-41.

Constantinou, P. (2017). Instructional assessment strategies for health and physical education. *Strategies: A Journal for Sport Educators, 30*(3), 3-9.

Creswell, J. (2003). *Research Design- Qualitative and Mixed Methods Approaches.* Lincoln: University of Nebraska,SAGE Publications.

Ćukušić, M., Željko, G., & Jadrić, M. (2014). Online self-assessment and students' success in higher education institutions. *Computers and Education, 72*(C), 100-109.

Darrow, A.-A., Johnson, C. M., Miller, A. M., & Williamson, P. (2002). Can students accurately assess themselves? Predictive validity of student self-reports. *Update Applications of Research in Music Education, 20*(2), 8-11.

DeLuca, C., & Klinger, D. (2010). Assessment literacy development: Identifiying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy and Practice, 17*, 419-438.

Denscombe, M. (2003). *The Good Research Guide For Small-scale Research Projects.* Maidenhead-Philadelphia: Open University Press.

Denscombe, M. (2007). *The Good Research Guide for small-scale social research projects.* New York: Mc Graw Hill.

Department of Education, a. t. (1996). *A Curriculum Framework for Physical Education:Adjusting the focus.* Newfoundland: Memorial University.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education. *Pracical assessment, research and evaluation., 14*(7), 1-11.

Earl, L. (2003). *Assessment as learning: using classroom assessmnet to maximise student learning.* Thousand Oaks, CA: Corwin Press.

Eberline, A. D., & Richards, K. R. (2013). Teaching with technology in physical education. *Strategies, 26*(6), 38-39.

Education Act. (1988). Malta. Retrieved from http://www.justiceservices.gov.mt/DownloadDocument.aspx?app=lom&itemid=8801

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550.

Elwood, J., & Klenowski, V. (2002). Creating communities of shared practice: the challenges of assessment use in learning and teaching. *Assessment & Evaluation in Higher Education, 27*, 243-256.

Employment, M. f. (2012). A National Curriculum Framework for All. Education Department, Malta.

Employment, M. f. (2015). *National Quality Assuarance Framework for Further and Higher Education.* Malta: Ministry for Education and Employment.

Employment, M. o. (2014). *Framework For the Education Strategy for Malta 2014-2024: Sustaining Foundations, Creating Alternatives, Increasing Employability.* Malta: Ministry of Education and Employment.

Falzon, J. (2016, May). Promoting secondary school studnents' lifelong engagement in physical activity: leadong change through quality physical education. pp. 1-122.

Frapwell, A. (2010). Assessment in physical education- fit for purpose? We've got an APP for that! *Physical Education Matters, 5*(3), 13-18.

Gibbons, S. L., & Kankkonen, B. (2011). Assessment as learning in physical education: making assessment meaningful for secondary school students. *Physical & Health Education Journal, 76*(4), 6-12.

Goldkuhl, G (2012) Pragmatism vs interpretivism in qualitative information systems research, *European Journal of Information Systems,* 21:2, 135-146

Gregory, K., Cameron, C., & Davies, A. (1997). *Setting and using criteria: foruse in middle school and secondary school classrooms.* Courtney: BC: Connections Publishing.

Grima, G., & Chetcuti, D. (2003). Current assessment practices in schools in Malta and Gozo- a research project. *Journal of Maltese Education Research, 1*(2), 57-94.

Handbook for the Teaching of Physical Education. (2012). Malta: Directorate for Quality and Standards in Education.

Hardman, K. (2008). The situation of physical education in schools: a european perspective. *Human Movement, 9*(1), 5-18.

Hardman, K. (2011). Global issues in the situation of physical education in schools. In K. Hardman, & K. Green, *Conemporary issues in Physical Education: International Perspectives* (pp. 11-28). Auckland: Meyer and Meyer Sport.

Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education, 35*, 231-224.

Harlen, W. (2004). *A systematic review of the evidence of the impact on students, teachers and the curriculum of the process of usiing assessment by teachers for summative purposes.* Institute of Education, Socila Science Unit. London: Evidence for Policy and Practice Information and Co-Ordinating Centre.

Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education, 10*(2), 169-207.

Hatfield, S., & Philips, R. (1989). *Records of Achievement in Physical Education, Topic No. 3.* NWCPEA.

Hattie, J. (2012). *Visible learning for teachers: Maximising impact on learning.* London: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hay, P., & Penney, D. (2009). Proposing coditions for assessment efficacy in physical education. *European Physical Education Review: 5*, 389-405.

Hay, P., & Penney, D. (2013). *Assessment in physical education: a sociocultural perspective.* London: Routledge.

Hay, P., & Penney, D. (2013). *Assessment in Physical Education: A sociocultural perspective.* Abingdon, Oxon: Routledge.

Heitink, M., Van der Kleij, F., Veldkamp, B., Schildkamp, K., & Kippers, W. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review, 17*, 50-62.

Hendriks, M., Scheerens, J., & Sleegers, P. (2014). Effects of evaluation and assessment on student achievement: a review and meta-analysis. In M. Hendriks (Ed.), *The influenve of school size, leadership, evaluation and time on student outcomes* (pp. 127-174). University of Twente: Enschede.

Hsieh, H., & Shannon, S. (2007). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288.

Huba, M., & Freed, J. (2000). *Learner-centered assessment on college campuses: shifting the focus from teaching to learning.* Boston: Allyn and Bacon.

Hussain, A., Said, A., & Hilal, A. (2014). Classroom assessment: teacher practices, student perceptions and academic self efficacy beliefs. *Social Behavior and Personality, 42*(5), 835-856.

Job Description Teacher. (n.d.). Retrieved from education.gov.mt: http://education.gov.mt/en/resources/Documents/Policy%20Documents/TeachersCodeOfEthics.pdf

Johnson, B., & Turner, L. (2003). Data collection strategies in mixed methods research. In A. Tashakkori, & C. Teddlie, *Handbook of mixed methods in social and behavioral research.* (pp. 297-319). Thousand Oaks, CA: Sage.

Johnson, R. (2001). Peer Assessments in Physical Education. *Journal of Physical Education, Recreation and Dance, 75*(8), 33-40.

Johnson, R. (2004). Peer assessment ain physical education. *Journal of Physical Education, Recreation and Dance, 75*(8), 33-40.

Johnson, S. (2013). On the reliability of high-stakes teacher assessment. *Research Papers in Education, 28*(1), 91-105.

Juniu, S. (2011). Pedeagogical uses of tehcnology in Physical Education. *Journal of Physical Edcucation, Recreation and Dance , 82*(9), 41.

Kinnunen, D., & Lewis, D. (2013). A case study of preservice physical education teachers' attitudes toward and percieved barriers to quality education. *Journal of Education and Practice, 4*(8), 123-133.

Klenowski, V. (2009). Assessment for learning revisited: an Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice, 16*, 263-268.

Kretschmann, R. (2015). Physical education teachers'subjective theories about integrating information and communication (ICT) into ohysical education. *The Turkish Online Journal of Educational Technology, 14*(1), 68-96.

Laker, A. (2001). *Developing Personal, Social and Moral Education through Physical Education.* New York: Routledge Falmer.

Lee, A., Keh, N., & Magill, R. (1993). Instructional effects of teacher feedback in physical education. *Journal of Teaching Physical Education, 12*(3), 228-243.

Lopez-Pastor, V., Kirk, D., Lorente-Catalan, D., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical Education: a review of international literature. *Sport, Education and Society, 18(1)*, 57-76.

Lund, J., & Kirk, M. (2002). *Performance-based assessment for middle and high school physical education.* Champaign: Hument Kinetics.

Lund, J., & Tannehill, D. (2010). *Standards-based Physical Education Curriculum Development.* Boston: Jones and Bartlett.

Marshall, C., & Rossman, G. (2016). *Designing Qualitative Research.* London: Sage Publications.

Mc Neill, P., & Chapman, S. (2005). *Research Methods.* London: Routledge.

McConnel, D. (2006). *e-learning groups and communities: Imagining learning at the age of the internet.* OU Press.

McDowell, L., Sambell, K., & Davison, G. (2009). Assessment for learning: a brief history and review of terminology. *Improving Student Learning Through the Curriculum* (pp. 56-64). Oxford: Oxford Centre for Staff and LEarning Development.

Mercier, K., & Dolittle, S. (2013). Assessing student achievement in physical education for teacher evaluation. *Journal of Physical Education, Recreation and Dance, 84*(3), 38-42.

Ministry of Education and Employment. (2012). *A National Curriculum Framework for All.* Salesian Press-Malta.

Mohan, R. (2016). *Measurement, Evaluation and Assessment in Education.* Delhi: PHI Learning.

Mohan, R. (2016). *Measurment, Evaluation and Assessment in Education.* Delhi: PHI Learning Private Limited.

Monsterrat, M. (2007). A Personal Narrative:"Self Assessment- A State of Utopia?". *British Educational Research Association Annual Conference.* London: Instute of Edication.

Moss, C. M. (2013). Research on Classroom Summative Assessment. In J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment* (pp. 235-255). Thousand Oaks: SAGE Publications, Inc.

Mosston, M., & Ashworth, S. (2002). *Teaching Physical Education 5th Edition.* Unitied States: Pearson Education.

National Minumum Curriculum. (2000).

Ndalichako, J. L. (2015). Secondary school teachers' perceptions of assessment. *International Journal of Information and Education Technology, 5*(5).

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

Nitko, A., & Brookhart, S. (2007). *Educational Assessemt of Students.* Upper Saddle River: NJ: Merril.

Nutly, D. (2008). The adequacy of response rates to online and paper surveys. *Assessment & Evaluation in Higher Education*, 301-314.

Onwuwgbuzie, A., Johnson, R., & Collins, K. M. (2011). Assessing Legitimation in mixed research: a new framework. *Qual Quant 45*, 1253-1271.

Pedder, D., & James, M. (2012). Professional learing as a condition for assessement for learning. In J. Gardner (Ed.), *Assessment and Learning* (pp. 33-48). London: Sage.

Penney, D., & Hay, P. (2008). Inclusivity and senior physical education: insights from queensland and western Australia. *Sport, Education and Society, 13*(4), 431-452.

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

Physical Education: Teaching Objectivesand Learning Outcomes. (2012). Malta: Directorate for Quality and Standards in Education.

Popham, W. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice, 48*, 4-11.

Puzanova, Z. V., Larina, T., & Zakharova, S. (2018). Recommendations for interviewers conducting sociological surveys: the use of pschotypes theory and analysis of respondents' nonverbal reactions. *RUDN Journal of Sociology, 18*(1), 156-165.

Redelius, K., & Hay, P. (2012). Student views on criterion-referenced assessment and grading in Swedish physical education. *Physical Education and Sport Pedagogy, 17*(2), 211-225.

Rink, J. (2010). *Teaching physical Education for learning .* New York: mcGraw-Hill.

Robson, C. (2011). *Real World Research.* Corwall: Wiley.

Rodriguez, M. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurment in Education, 17*(1), 1-24.

Sach, E. (2012). Teachers and testing: an investigation into teachers' persceptions of formative assessment. *Educational Studies, 38*(3), 261-276.

Sadler, R. (1989). Fromative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Salkind, N. J. (2010). Cross-Sectional Design. In N. J. Salkind, *Encyclopedia of Research Design* (p. 314). Thousand Oaks: SAGE Publications, Inc.

Schleicher, A. (2018). *World class: How to build a 21st-century school system, strong performers and successful reformers in education.* Paris: OECD Publishing.

Shepard, L., Hammerness, K., Darling-Hammond, L., & Rust, F. (2005). Assessment. In L. Darling-Hammond, & J. Brandsford, *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 275-326). San Francisco, CA: Jossey-Bass.

SPARK. (2016, January 12th). *Using of Technology in Physical Education*. Retrieved from Spark PE: https://sparkpe.org/blog/using-technology-in-physical-education/

Stobart, G. (2008). *Testing times: The uses and abuses of assessment.* Abingdon, England: Routledge.

Sturgis, P., Smith, P., & Hughes, G. (2006). *A study of suitable methods for raising response rates in school surveys.* London: Department for Education and Skills.

Sturman, A. (1999). Case study methods. In J. Keeves, & G. (. Lakomski, *Issues in Educational Research* (pp. 103-112). Oxford: Elsevier Science.

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

Thorburn, M. (2007). Achieving conceptual and curriculum coherence in high-stakes school examinations in physical education. *Physical Education and Sport Pedagogy, 12*, 163-184.

Togfors, B., & Öhman, M. (2016). The implications of assessment for learning in physical education and health. *European Physical Eduation Review*, 150-166.

Tolgfors, B. (2018). Different versions of assessmnet for learning in the subject of physical education. *Physical Education and Sport Pedagogy*, 310-327.

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom using action research to explore and modify theory. *British Educational Research Journal, 27 (5)*, 615-631.

Trauth-Nare, A., & Buck, G. (2011). Assessment for learning. *The Science Teacher, 78*(1), 34-39.

Turner, D. W. (2010). Qualitative Interview Design: A Practical Guide for Novice Investigators. *The Qualitative Report 15:3*, 754-760.

Van Mol, C. (2017). Improving web survey efficiency: the impact of an extra reminder and reminder content on web survey response. *International Journal of Social Research Methodology*, 317-327.

Van Selm, M., & Jankowski, N. W. (2006). Conducting online surveys. *Quality and Quantity, 40*(3), 435-456.

Van-Vuuren Cassar, G. (2011). Assessments and qualifications in physical education. In K. Hardman, & G. K, *Conemporary issues in Physical Education: International Perspectives* (pp. 194-214). Auckland: Meyer and Meyer Sport.

Veal, M. L. (1988). Pupil assessment issues: A teacher educator's perspective. *Quest, 40*, 151-161.

Von Diether, B. (2016, December 5th). *Writing Chapter 3: The Methodology*. Retrieved June 6th, 2018, from Dissertation Writing: https://dissertationwriting.com/2016/12/05/write-dissertation-methodology-help/

Wiliam, D. (2011). *Embedded formative assessment.* Bloomington: Solution Tree press.

William, D. (2006). Assessmet for learning: why,what and how. *Excellence in Assessment: Assessment for Learning* (pp. 2-16). Cambridge: Cambridge Assessment.

William, D., & Thompson, M. (2007, April). Tight but Loose: A Conceptual Framework for Scaling Up School Reforms. Chicago, IL: Americal Educational Research Association (AERA).

Wyayy-Smith, C., Klenowski, V., & Colbert, P. (2014). Assessment understood as enabling. In C. Wyayy-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 1-20). Dordrecht, The Netherlands: Springer International.

Yin, R. K. (1994). *Case Study Research: Design and Methods.* London: Sage Publications.

Zigo, D., & Moore, M. T. (2002). English Education. *Social Science Premium Collection, 34*(2), 137.