



Uio • University of Oslo

Learning-oriented Assessment in Norwegian Physical Education Classes

The Teacher's Perspective

Thomas Sandholt Lund

Master of Philosophy in Comparative and International Education

Department of Education

Faculty of Educational Sciences

University of Oslo

December 2020

Abstract

While learning-oriented assessment (LOA) is popular in educational research and Norway has adopted assessment for learning (AFL) to its regulations, research suggest that Norwegian physical education (PE) is out of line with principles about LOA. This thesis has two explorative purposes, firstly, it explores assessment practices used by Norwegian PE teachers and evaluates these practices in relation to LOA. Secondly, it investigates the barriers to implementing LOA practices for teachers. Qualitative semi-structured interviews were conducted with 12 physical education teachers from 10 Norwegian secondary schools, two of which were private schools.

The findings of this thesis show that teachers use both testing and authentic assessment methods in PE. However, use of testing for grading purposes is not evident, but is used for formative purposes. Observation and video constitute most of the authentic assessment. Teachers attempt to provide formative feedback to pupils, but the degree to which they manage to do this vary. Self-assessment is used on an occasional basis, more as a tool of ‘reality-orientation’ rather than for the engagement of pupils as owners of their own learning. When it comes to the challenges to LOA, it appears that time to see and assess pupils represent the major barrier to the implementation of LOA. Teachers feel they do not have enough time to see pupils and not enough time to provide continuous and quality feedback. The pressure to document performances seems to be one major driver to this lack of time, alongside short sessions and big class sizes. Challenges also seem to lie within teacher education, resources for teachers, unmotivated pupils and facilities.

Based on these findings, I argue that Norwegian PE teachers struggle with implementing LOA and both the time aspect, documentation needs, few sessions and big classes represent barriers to LOA. Teachers seem to lack clear methodology that can help them with the implementation of AFL strategies. This thesis helps reducing the knowledge gap about challenges to implement LOA in PE. Much research confirms promising results of formative assessment and much research highlights problematic sides of assessment in PE, while little research deals with LOA in PE. This thesis contributes to knowledge about why PE teachers struggle with implementing LOA.

Keywords: Physical Education, Learning-oriented Assessment, Assessment for Learning, Assessment Practices, Assessment Methods, Assessment Challenges

Acknowledgements

This thesis has been a long, and learning-oriented journey which I am very thankful of. 2020 has been a weird and challenging year and I am very grateful for the patience and support of both my girlfriend and family. I could not conduct this thesis without their support. I also thank my employers and colleagues at the University of Western Australia, they have shown great patience during periods where this thesis has required all of my attention.

I would like to thank my supervisor Rogers Kaliisa, who has provided me with vital support and advice throughout a research process that has been new and challenging to me. He has been supportive and solution-oriented in a year that has been challenging both in terms of the pandemic, personal health, work obligations and time devoted to my passion, which is sport. I am forever grateful for the support I received from Rogers when I wanted to change my thesis topic to a sport-related topic, which has become an immensely educational experience for me, which I can benefit from in my roles in the sports and health industry.

I would like to thank everyone in my class in the CIE program, who has inspired me with their competence, dedication and perspectives. I have learnt so much from this class and hope everyone will succeed in personal and professional careers. I thank the course coordinators and all lecturers for great learning experiences the last couple of years!

Table of Contents

Abstract.....	iii
Acknowledgements.....	iv
Table of Contents.....	v
List of Abbreviations.....	vii
List of Tables.....	viii
1 Introduction.....	1
1.1 Background of Physical Education in Norway.....	2
1.2 Focus and Purpose of this Thesis.....	3
1.3 Research Questions.....	5
1.4 Clarification of key terms.....	5
1.5 Limitations.....	7
1.6 Structure of the Thesis.....	8
2 Theoretical perspective and literature review.....	10
2.1 Theoretical perspective.....	10
2.2 Literature review.....	13
2.2.1 Research context.....	14
2.2.2 Research question 1: assessment practices.....	23
2.2.3 Research question 2: challenges to LOA.....	32
3 Methodology.....	43
3.1 Research design.....	43
3.2 The sampling process.....	43
3.3 Research interviews.....	45
3.3.1 Qualitative semi-structured interviews.....	45
3.3.2 Interviews.....	46
3.4 Transcription.....	47
3.5 Research analysis.....	48
3.6 Trustworthiness and authenticity.....	49
3.6.1 Trustworthiness.....	49
3.6.2 Authenticity.....	50
3.7 Ethical considerations.....	51
4 Presentation of Findings and Discussion.....	54
4.1 Presentation of the teachers.....	54
4.2 Research question 1: What forms of assessment practices do PE teachers use, and to what extent do they align with learning-oriented assessment practices?.....	56

4.2.1 Clarifying and sharing learning intentions and criteria for success	57
4.2.2 Engineering effective classroom discussions, questions, and learning tasks that elicit evidence of learning	58
4.2.3 Providing feedback that moves learners forward.....	66
4.2.4 Activating pupils as instructional resources for one another	67
4.2.5 Activating pupils as the owners of their own learning.....	68
4.3 Research question 2: are there barriers to teachers using learning-oriented assessment practices in their PE practice, and what kinds of support or advice do physical education teachers need to develop educationally sound, successful and sustainable forms of assessment?	69
4.3.1 Teachers' view on learning in PE	71
4.3.2 Student and parent expectations.....	73
4.3.3 Documentation needs	74
4.3.4 Motivation and stress among the pupils.....	76
4.3.5 Pupils' (lack of) participation	77
4.3.6 Competence goals and local learning plans	79
4.3.7 Teacher Education.....	80
4.3.8 Resources for teachers	81
4.3.9 Class size and time to assess	83
4.3.10 Facilities and resources	85
4.3.11 The biggest challenge.....	86
4.4 Private vs public schools.....	88
5 Conclusion	91
5.1 Research question 1: what forms of assessment practices do PE teachers use, and to what extent do they align with learning-oriented assessment practices?	91
5.2 Research question 2: are there barriers to teachers using learning-oriented assessment practices in their PE practice, and what kinds of support or advice do physical education teachers need to develop educationally sound, successful and sustainable forms of assessment?	93
5.3 Private vs public schools.....	94
5.4 Contribution to the field.....	95
5.5 Future research.....	95
Reference List	97
Attachments	105
Attachment 1:.....	106
Attachment 2:.....	109
Attachment 3:.....	112

List of Abbreviations

AFL	Assessment for learning
ARG	Assessment Reform Group
CAQDAS	Computer-assisted qualitative data analysis software
LOA	Learning-oriented assessment
MBO	Management by objectives
NPM	New public management
PE	Physical education
PFT	Physical fitness test
SDT	Self-determination theory
UDIR	Norwegian Directorate for Education and Training

List of Tables

Table 1: formative assessment strategies by William & Thompson (2007).....	13
Table 2: showing the demographic characteristics of the participating teachers.....	54
Table 3: themes and sample extracts for research question 1.	56
Table 4: themes and sample extracts for research question 2.	69

1 Introduction

In a world where health and education have become important factors for governments to pay attention to, physical education (PE) represent a meeting point between these two sectors. While ‘teaching’ pupils to be physically active necessitates the facilitation of intrinsic motivation (Ryan & Deci, 2020), the educational context of PE necessitates evaluation of pupils learning by a teacher, which represents an extrinsic motivation for pupils. While assessment serves an important summative role of evaluating competence and skill levels in both countries and pupils, research shows that assessment also has an important role in the learning process (Gardner, 2012). Formative assessment (Black & William, 2009) and assessment for learning (AFL) are associated with what in this thesis will be called ‘learning-oriented assessment’ (LOA). More than merely evaluating skills and abilities, LOA also facilitates the learning process for pupils.

In Norway, the goal of PE is not only to increase the competence and skills of pupils, but to inspire them to lifelong physical activity and thus facilitate for good long-term national health (Utdanningsdirektoratet, 2015). Research show that pupils who experience PE as consistent with learning and personal development report higher enjoyment with the subject (Jaakkola et al, 2016). Also, young adolescents who have good perceptions about their physical abilities are more likely to be involved in physical activity in the long term (Jaakkola et al, 2015), in line with self-determination theory (Ryan & Deci, 2020). Self-determination theory (SDT) stresses the importance of autonomy for motivation, where autonomy fits with formative assessment’s goal of making pupils into owners of their own learning, rather than merely being subjects to summative evaluation (Black & William, 2009). Competence is another aspect of SDT, which stresses the need for a focus on learning in PE. Research suggests that formative assessment has positive motivational outcomes in the classroom seen through the lens of the self-determination theory (Harlen, 2012; Hondrich et al, 2018).

With regards to SDT, assessment must be said to be a key aspect on how pupils perceive their own competence and abilities and ultimately how they are motivated in the subject. Thus, in addition to facilitating learning in PE, assessment also represent a component that either work in favour of or against goals of PE such as promoting lifelong physical activity. While the Norwegian government wants PE to be a facilitator for lifelong physical activity, research suggests that PE in Norway are more centred around summative rather than formative assessment (Græsholt, 2011; Arnesen, Nilsen & Leirhaug, 2013; Leirhaug & MacPhail,

2015) with implications for LOA. This thesis will look at the assessment practices of PE teachers in Norway, as well as the challenges to implementing LOA strategies (Black & William, 2009) in Norway seen through the eyes of the teachers.

1.1 Background of Physical Education in Norway

PE in Norway can be described as a subject concerned with health and the official learning plan highlights a physically active lifestyle and lifelong activity enjoyment as the main purposes of the subject (Utdanningsdirektoratet, 2015). Pupils are supposed to learn movement cultures, including swimming and outdoor recreation and the social aspect of the subject is supposed to be an arena for learning fair play and respect for others. Through different movement cultures, sports, fair play, dance, swimming and outdoor recreation and a focus on health, nutrition and exercise, pupils are supposed to experience joy, mastery and inspiration that inspire to lifelong physical activity and mastery of own bodies. If the assessment methods in Norwegian PE classes do not work in the favour of learning and thus mastery, one could ask how well aligned assessment is with the purposes of the subject.

The previous education reform in Norway, Knowledge Promotion in 2006, was a result of poor performance in international educational surveys such as PISA. The reform was a part of a wider international focus, new public management (NPM), which made schools and teachers more accountable for pupil learning, to increase national testing and to manage education by national objectives rather than by rules or principles (Møller & Skedsmo, 2013, p. 347). The idea is to make government branches work more like private companies. Despite the neoliberal context and the accountability objectives, the 2006 curriculum also brought ideas from Black and William's work on formative assessment (1998; 2009) and from the Assessment Reform Group (ARG), which have done much of the leading work on AFL (Gardner, 2012). The curriculum introduced some features inspired by work on AFL, such as sharing assessment criteria with the pupils, facilitating for self-assessment and giving them half-yearly assessments (Leirhaug & Annerstedt, 2016, p. 617). However, research suggests that the introduction of AFL in Norway has not been successful, and that there are great challenges when it comes to the implementation of such practices (Leirhaug & Annerstedt, 2016). This creates the backdrop for this thesis, where I will investigate the challenges to implementing assessment practices that focuses on learning rather than evaluation.

The international context of education, as well as assessment in PE can be associated with neoliberal ideas about accountability and efficiency, and strategies such as high stakes testing and physical fitness testing (PFT). Macdonald (2011) points to PE as an interesting point of analysis when it comes to neoliberalism, as a subject in the intersection between education and health agendas (p. 38). Møller and Skedsmo (2013) points to that due to the strong emphasis on equity and social democratic values, elements of NPM are not as central in Norway as elsewhere. Marketisation and competition remain limited in the Norwegian education system, but some privatisation has occurred in the school system and management by objectives (MBO) has become a part of the accountability of the education sector. The prominence of testing has grown, the accountability of teachers and simultaneously, focus on the rights of pupils has grown (p. 348). This means that the growing trend of testing pupils, is accompanied by their rights to have their learning performance documented, which adds to the complexity of the assessment process for teachers.

PE in Norway is not under a national testing regime or obligatory testing (Utdanningsdirektoratet, 2015). The assessment system is decentralised and decided by the teacher, who designs the assessment in line with the competence goals in the learning plan. Testing in PE is regardless of this still linked to testing, or the lack of LOA in Norway (Græsholt, 2011; Arnesen, Nilsen & Leirhaug, 2013, p. 25). Although an alignment between learning theory and assessments is a prerequisite for the validity of assessments, there has traditionally been a misalignment between the development of pedagogy and classroom assessment (James & Lewis, 2012, p. 189). With the introduction of AFL in PE in Norway as well as aims to inspire to lifelong physical activity, assessment represents a major challenge for PE in terms of its validity and in terms of promoting learning.

1.2 Focus and Purpose of this Thesis

This thesis will focus on LOA methods in Norwegian PE classes during secondary education (year 10-13). In his doctoral thesis, Leirhaug points to (2016, p. 3) that although there is an increasing amount of literature about assessment in PE, there is little research that focuses on formative assessment practices. With ever-growing interest in LOA and the interest of the Norwegian government in implementing this (Utdanningsdirektoratet, 2018), the overarching purpose of this thesis is to investigate how Norwegian PE teachers are coping with LOA and what challenges may be.

The thesis has two explorative purposes, the first relating to the assessment practices used by Norwegian PE teachers, while the other is to explore the challenges teachers meet when it comes to assessment in the subject from a learning-oriented perspective. The first purpose is important to understand what practices and strategies teachers use in order to evaluate how these practices relate to literature on ‘good practice’ when it comes to assessment practices that promotes learning. Without a clear understanding of the teachers’ practices, it will be impossible to analyse how these can become more ‘learning-oriented’.

The second purpose of the thesis aims to explore the challenges PE teachers face when it comes to assessment, which may problematise LOA. For example, if the teacher feels that he or she does not have enough time for LOA assessment, this might lead to the teacher undertaking testing strategies, where pupils are given activities that are only valuable for assessment of learning, rather than assessment for learning. Lopez-Pastor and colleagues’ review of ‘alternative’ assessment methods in PE (2012, p. 71) asks what the barriers to implementing ‘alternative assessment’ in PE are, as opposed to traditional methods of assessments, such as PFTs. In a study of assessment in a Norwegian PE class, testing represented the dominant assessment method and pupils had no experience with ‘other’ methods (Græsholt, 2011, p. 108), which supports the notion of traditional assessment methods, as opposed to ‘alternative’ methods. The question is whether this is still valid and whether the results will reveal evidence of LOA.

Secondary education is the level of study which gets the most PE classes per year in the Norway (Utdanningsdirektoratet, 2015), which makes it an interesting level to look at in this thesis. Secondary education is also a level of study where one can find an increasing drop-out rate from organised sports (Bakken, 2019, p. 73; Lagestad & Sørensen, 2018, p. 94), and while I do not know about the workings of this, nor attempts to do so in this thesis, it increases the importance of PE as an arena to learn to enjoy physical activity. In other words, the official learning plan’s goal of stimulating to lifelong enjoyment with physical activity finds itself in a crucial phase during secondary education. Investigating the effects of assessment practice on pupils’ lifelong physical activity levels would be very challenging and is way beyond the scope of this master’s thesis. However, exploring the relationship between assessment practice and learning in PE can at least give an idea about whether the current structures of PE allows pupils to be equipped with the tools and capacities to engage with physical activity, which is set out to be the purpose of the subject.

The study will look at the teacher's perspective, in order to get a greater understanding of the opportunities and challenges PE teachers face when it comes to LOA. While both the pupil-perspective, as well as the school leader and policy-makers perspective is interesting when it comes to the question of LOA in Norwegian schools, this study will focus at the person in charge of designing and implementing the assessment.

1.3 Research Questions

1. What forms of assessment practices do PE teachers use, and to what extent do they align with learning-oriented assessment practices?

This first research question will work to uncover the assessment practices commonly used by Norwegian PE teachers and evaluate how these relate to LOA strategies, in this case, represented by Black & William's framework on formative assessment. This research question may confirm what other research suggests, that Norwegian PE assessment practices are out of line with LOA (Arnesen, Nilsen & Leirhaug, 2013, p. 25; Græsholt, 2011), or it may expose evidence of AFL, which has been an official focus area of the government (Utdanningsdirektoratet, 2018).

2. Are there barriers to teachers using learning-oriented assessment practices in their PE practice, and what kinds of support or advice do physical education teachers need to develop educationally sound, successful and sustainable forms of assessment?

The second research question will investigate challenges or barriers that teachers face in the implementation of LOA in Norwegian PE. Are there issues related to teacher education, time constraints or learning plan barriers that work against LOA? Are there other issues, or does the implementation stop with the teachers? This research question will contribute to closing some of the research gaps in the challenges to implementing assessment practices that promote pupil learning (Lopez-Pastor et al, 2012, p. 71).

1.4 Clarification of key terms

Learning-oriented assessment (LOA) – is a concept that appear many places in literature and that may have many different definitions. In this thesis, LOA represent a less strict concept than both formative assessment and AFL. The meaning aligns with the definition proposed by Cambridge Assessment (n.d.), that LOA deals with assessment on multiple levels and that its

intention is to change assessment's traditional relationship to learning, making them more interconnected. While research suggests assessment in PE is not fully aligned with theory about formative assessment (Black & William, 2009), LOA in this thesis represent a broader concept that emphasise assessment with the intention of creating improved learning situations for pupils. The key point is learning as the objective.

Formative assessment – in this thesis is aligned with the definition by Black and William, which forms the basis for their theoretical framework about formative assessment. Formative assessment is defined by using information about pupil achievement to improve instruction (2009, p. 9). By gathering information about pupil knowledge through assessment, the teacher can make better decisions about the next steps in the learning process for the pupils. According to Gardner (2012, p. 2), formative assessment dates to Scriven in 1967, and thereby is the older concept of the two concepts formative assessment and AFL.

Assessment for learning – is different from formative assessment in the way that while formative assessment originally could be interpreted as a form of summative assessment that happened continuously but that was not necessarily used to improve learning, AFL is unambiguous about having learning as the main motivation (Gardner, 2012, p. 2). Later both AFL and formative assessment are concepts linked to learning, but formative assessment may be used in a summative setting, where it is not used to improve learning.

Authentic assessment – represent an assessment method where the teacher let the pupil conduct the activity he or she is supposed to be assessed in, whereby the teacher evaluates the pupil's ability through observation (Engvik, 2010, p. 209).

Alternative assessment – represents all methods of assessment that goes beyond traditional assessment methods such as testing (Lopez-Pastor et al, 2012).

Traditional assessment – assessment that is linked to behaviouristic ideas about the acquisition of skills. It attempts to test competence at single points in time, such as in a PFT or a 'skill track'.

Learning plan – is directly translated from Norwegian and while it can be referred to as a curriculum, learning plans are made for individual subjects and are updated more often than the national curriculum. The current learning plan in PE is from 2015 and the learning plans for PE have been adjusted a few times since the introduction of the latest national curriculum.

Competence goals – are the learning goals in the learning plan. These are broad and guide PE teachers in the formation of teaching and assessment strategies. After tenth grade, the pupils are for example supposed to practice fair play, navigate by using map and compass and explain the relationship between physical activity, lifestyle and health (Utdanningsdirektoratet, 2015). There are 15 competence goals that apply to PE in secondary school and that work as reference points for PE teachers when they assess pupils.

1.5 Limitations

As important as what this study's research question will answer, is what it cannot answer. Limitations reside both in the research questions, the scope of the research and the methodology of the research. One limitation of the research lies in the methodology, as qualitative research interviews with PE teachers is the only method of this study. While I argue that is the best approach to answer my research questions, more insights to the topic could be drawn from both pupil interviews, ethnography and document analysis in conjunction with teacher interviews. However, this is beyond the scope of a master's thesis as the material would have grown too massive for a year-long project for a researcher without previous knowledge on the subject area. The semi-structured interviews were broad and allowed me to identify key areas related to my research and to get an overview of the topic as a less experienced researcher. On the other hand, this broad approach and the scope of this thesis obviously limits the possibility to dig deeper into interesting questions.

One limitation of this thesis specifically, is that I could align the interview guide with the framework more, to get more detailed responses for each of the formative strategies. For the first formative strategy by William and Thompson, I lack clear responses from this from all participants. The fact that I did not ask specifically for AFL or LOA in the interview guides can be argued to be a weakness, as I may miss out on thoughts about this from some teachers. On the other hand, I argue that this has also been a strength in that my questions have not been leading questions in this sense. I am glad that evidence regarding the assessment practices and challenges to LOA have come out of the teachers without me putting concepts such as AFL or LOA in their mouths. I argue that if they do not mention such concepts themselves throughout the close to hour-long interviews, the concepts are not as entrenched in the practices as they should be.

I argue that despite potential limitations with teachers' biases, or unwillingness to expose less good habits when it comes to assessment and the lack of the pupil perspective, teachers are the best reference point for this study. Answering my research questions would not be possible without interviews with teachers. Another limitation is the geographical context of this study. The findings of this study cannot, without carefulness and reflections about contextual factors, be found true in other countries, or contexts. Both historical factors, laws and policies, teacher education, nature of schools and classrooms as well as sociological factors may create completely different context in other contexts. What this study can do is to highlight PE teachers' assessment practices in the Norwegian context and reveal what challenges teachers face when it comes to the implementation of LOA in Norway.

1.6 Structure of the Thesis

This thesis is organised into six chapters. The first chapter sets the context and the background for why this particular research topic was chosen, including the historical background of PE, the context of assessment and the relationship between assessment and learning. It also narrows it down to the focus and purpose of this thesis, as well as setting the research questions. The idea of the chapter is to provide a narrative to why LOA in PE is worth looking into.

Chapter two provides the theoretical foundations for the research, with an overview of a framework on formative assessment by Black and William (2009), and William and Thompson (2007), which sets the guide for how to analyse the data with regards to formative assessment practices. The chapter also consists of a literature review on research related to assessment, learning, PE and challenges to assessment and learning in PE.

Chapter three discusses methodology, covering issues such as research strategy, research design, data collection, data analysis, trustworthiness and authenticity, as well as ethical considerations.

Chapter four will present and discuss the research's finding in the light of the theoretical framework by Black and William and relevant literature, where statements by the individual teachers will illuminate each of the five formative assessment strategies.

Chapter five will be the conclusion of this thesis, with a summary of the main findings in response to the two research questions, a comparison between private and public school teacher responses, contributions to the field, as well as recommendations for further research.

2 Theoretical perspective and literature review

2.1 Theoretical perspective

If this thesis is going to make sense in terms of adding value to the field, it needs a theoretical framework that confines it, which puts important concepts in a theoretical perspective.

Without a theoretical framework with a basis in existing literature, this thesis will have limited value when comparing it to other works on the topic. If the findings of this thesis are not properly connected to existing and important literature of learning-oriented, or formative assessment, it may be impossible to fully understand the findings in the vast literature on assessment. I will use Black and William's framework on formative assessment (2009) for exploring my research questions, which is based on William and Thompson's five formative strategies (2007). These strategies form the basis of formative assessment, that helps this thesis in connecting the research data to theory about LOA. In the first research question, the framework will help identifying which assessment practices have LOA elements and which assessment practices that do not. The second research question deals with challenges to implementing LOA strategies and it would be impossible to do this, without a clear understanding of what constitutes as LOA.

The potential learning gains of formative assessment is well-known through the literature of Black and William (1998), who through their meta-analysis of about 250 articles on formative assessment found strong positive impact on learning outcomes when using formative practices such as feedback. In 2009, they formulated a theoretical framework for formative assessment that guides this thesis. When looking into literature about assessment and learning, one cannot get around the work done by the ARG. This was a group of influential scientists on the topic who has gathered research on assessment and learning, to improve understanding of the role between these two factors that so often become conflicting interests (Gardner et al, 2012, p. 1). Black and William's article is heavily cited and builds on earlier contributions by them and others on assessment and learning, to create a unifying framework. Their work builds on their literature review from 1998 and on the work by William and Thompson (2007). Their framework also builds on criticism from Perrenoud (1998) to their 1998 review. This criticism deals with the lack of a theoretical foundation and attempts to unify and locate theory about formative assessment in relation to other pedagogic theories, as well as learning theory.

When looking at a theory about formative assessment, it is crucial to establish what is meant by ‘formative assessment’. Black and William suggests:

“Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited” (Black & William 2009, p. 9).

Black and William’s definition of formative learning focuses on whether information about pupil achievement is utilised by the agents of learning (teachers, peers, pupils) in order to improve instruction in the future. This is a simple and understandable definition that limits formative assessment to the use of information to propel the pupil forward, which will be used to define formative assessment in this thesis. ‘Classroom’ can easily be swapped with ‘sport field’ or ‘gymnasium’, without hampering the relevancy of the definition.

Black and William’s article draws on William and Thompson’s article (2007) which developed the five strategies for formative assessment. These strategies locate in a model that crosses the three different subjects (teacher, peer and learner) with three stages of the learning process, developed by Ramprasad (1983). These phases are (1) where the learner is going, (2) where the learner is currently and (3) how the learner can get there. In other words, the model aims to establish what the learning objectives require from the pupil, the pupil’s current level of abilities and what strategies that are necessary to get the pupil to the intended learning objectives. The model describes these strategies at three different levels of agents, from the teacher to peers to the learner. I will briefly describe these strategies.

The first strategy is to clarify the learning intentions and criteria for success, which feeds into the first stage of the process. This occurs at all three levels, from teacher to peer to learner. This can take the form of the teacher lecturing the pupils about the learning objectives and what is required from them, whereby the peers and the learner discuss this in class and help each other understanding the criteria. The strategy guides the data collection process by highlighting the need to ask about the teachers’ sharing of the learning criteria. Do teachers spend time enough to explain the learning objectives to pupils and letting pupils know what is expected by them? This has been highlighted in the research interviews.

The second strategy occurs at the second stage and is to engineer effective class-room discussions and tasks that reveal the pupils' current level of knowledge and understanding. This strategy occurs only at the teacher level. In PE, this can be a situation where the teacher sets up a game or play situation, for example in volleyball, both to see where the pupils are at in terms of game understanding, technique, team play and fair play. The activity should reveal where the pupils are in the learning process and what needs to be done. This creates an important dimension for the analysis of the research interviews: are the teachers interacting with the pupils throughout the classes? Or do they sit on the stands, taking notes, without the opportunity to 'coach' pupils? This strategy is linked to interview questions about the methods teachers undertake throughout assessment.

The third strategy is also exclusively located at the teacher level and is to provide feedback to pupils so that they move forward in the learning process. In PE, this can be taking the pupil out of the activity to provide feedback in terms of technical, tactical or social aspects of the game, that guides the pupil towards the intended learning objectives. This is one of the crucial aspects of the study. Are teachers giving pupils feedback that aids learning, or are they merely getting some feedback in the mandatory half-yearly assessment about what grade they are heading towards? How often and how teachers give feedback are included in the interview guide.

The fourth strategy entails activating pupils as instructional resources for each other. This strategy is located at the second and third stage of the teaching-learning process, across the peer level of interaction. This means that peers who has mastered an activity should get the opportunity to act as role models and show them to the learner. This can for example be done in gymnastics, where the teacher lets a pupil with good technical abilities show and instruct others how to do the activity. This is a part of the shifting of responsibility of learning from the teacher to the learners, which we will come back to in the fifth strategy. Are teachers in the study taking advantage of pupils with advanced competence? Are pupils given the opportunity to help each other and to become active participants in the learning process, rather than recipients of the teacher's instructions? In the interview guide, there is explicit reference to peer assessment and whether this is used by the teachers.

The fifth strategy is to activate pupils as owners of their own learning. This occurs in the second and third phase of the teaching-learning process and at the learner's level. While the traditional view is to see the teacher as responsible for the learning process (Black & William

2009, p.7), this theory of formative education indicates that learners are be responsible for their own learning, while the teacher’s role is to facilitate and guide the learner in the process towards the intended learning objectives. The question is whether the teacher actually uses self-assessment as part of the assessment and whether this gives pupils a proper opportunity to reflect around their learning, or whether it’s merely a way of letting pupils guess what grade they will get. Questions about self-assessment are included in the interviews. In this thesis, it is necessary to question whether this shift of responsibility is apparent in PE classes, or whether classes are more instructional in nature.

Table 1: formative assessment strategies by William & Thompson (2007).

	Where the learner is going	Where the learner is right now	How to get there
Teacher	1 Clarifying learning intentions and criteria for success	2 Engineering effective class-room discussions and other learning tasks that elicit evidence of student understanding	3 Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	4 Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	5 Activating students as the owners of their own learning	

2.2 Literature review

The literature review is important both to establish what we already know about the research topic, but also to contextualise this research within the field and to link findings to other research (Bryman, 2016, p. 94). It is important to identify whether current research is relevant, whether it is controversial and if there are unanswered research questions. Literature reviews can be divided into two categories, narrative and systematic reviews. While systematic often are stand-alone reviews of literature, a narrative review more often works to build up a research paper (p. 91). This review belongs to the latter category, as it aims to illuminate a research topic, so that a proper analysis and discussion can be made in relation to this thesis’ research questions. The literature review needs to illuminate opposite sides of the topic, in order to understand the challenges with each approach (p. 92). In the case of the first research question in this thesis, illuminating opposite ends would be at one point research related to traditional approaches such as testing in PE, and at the other end of the spectrum, research related to formative, or learning-oriented, assessment practices on the other side.

The second research question seeks to explore the challenges related to the implementation of LOA practices, whereby is it natural to go into research about assessment in PE that goes into such challenges.

The literature review in this thesis have been conducted through a few methods. The first method that brought me to relevant data was searching through relevant data at Oria, a Norwegian online portal for research and scholarly works, accessible to Norwegian pupils. Key words related to the thesis' area of focus were used in a variety of combinations, such as 'assessment', 'formative assessment', 'assessment for learning', 'physical education', 'testing', and so on. The second method I have used is the so-called snowballing method, where I have looked up references used in literature obtained from Oria or elsewhere. This has been an iterative process, gathering references from other sources, up to a point where a sense of satiation is obtained, that is, when few new influential sources have emerged. This includes looking up authors that are well-known in the topic, which I determine through frequent sightings of an author's name in relevant literature. A final method or acknowledgement is the relevant literature I have obtained from my supervisor, who has provided me with some relevant literature that have informed my research.

This literature review consists of three sections. The first section goes into the research context, with areas such as the historical development of PE, the learning plan in PE, the decentralised system of teaching and assessment in Norway and finally looking into a government report about formative assessment in Norway. In the second section, I will investigate research about the first research question for this thesis and attempt to illuminate assessment practices in PE. In the third and final section of this literature review, I will look at research on the second research question, about challenges related to assessment. While the research question is linked to challenges to implementing LOA, the literature review needs to go broader than this, looking at issues related to assessment in PE in general, in order to get an understanding of potential challenges.

2.2.1 Research context

2.2.1.1 History of PE in Norway

To get an understanding and overview of what is valued in PE, it is useful to have a look at the historical development of PE in Norway, as well as what David Kirk names 'the idea of the idea of physical education' (2010), which has developed over time and gone through

different phases and meanings. Kirk's concept deals with what society thinks of as the idea of PE, including both government agencies, schools and teachers, thereby having an impact on the context PE resides in.

PE in Norway started in the nineteenth century, first with the introduction of voluntary gymnastics in 1848, but both a lack of funding, teachers and interest among politicians led to the subject not being prioritised in Norwegian schools (Augestad, 2003, p. 65). In 1889, the subject was made mandatory for city schools, while it still was voluntary on countryside schools. The rising interest was both due to military considerations with rising tensions against the Swedish union, as well as a mission to strengthen the spirit, develop discipline and strengthen the body, all of which were influenced by ideals from Greek and Roman antiquity. The Swedish type gymnastics implemented in Norway at the end of the nineteenth century was intended to make strong, powerful and disciplined individuals in an attempt by the Norwegian state to prepare the nation militarily. David Kirk describes the period as 'physical-education-as-gymnastics' (Kirk, 2010). A strong focus on discipline was not only for wider societal needs but as important for formation purposes and as a prerequisite for learning (Augestad 2003, p.64). In a school system influenced by behaviourist ideas about learning, discipline created the basis for transfer of learning from the teacher's instructions to the obedient pupil.

While late nineteenth century and early twentieth century PE had a purpose of disciplining and preparing the nation militarily, this focus lost momentum throughout the twentieth century. In the coming decades, PE developed into what Kirk describes as 'physical-education-as sport-techniques' (2010), where the focus shifted from strengthening and disciplining pupils through gymnastics in particular, to teaching them as many sports and sporting techniques as possible throughout the school year. While being a much more varied and 'play-oriented' than gymnastics, the PE-as-sporting-techniques paradigm of PE is described by Kirk as 'molecularised'. It is designed to teach and evaluate certain features of what the subject consists of, such as measurable or testable skillsets related to sporting activities, but not so much non-technical factors from the current learning plan (Utdanningsdirektoratet, 2015) such as tactical skills, fair play, cooperative skills and effort. Assessment was not a concern in this period, as it was easy for the teacher to observe those who were good in sports and those that were not. However, in this period, USA, Great Britain and Australia would start focusing on 'objective' testing between the 60s and the 80s, which

attempted to identify motor skills and fitness, which is still a common assessment practice today (Lopez-Pastor et al, 2012, p. 58).

As a more modern approach, designed to teach pupils a more varied and relevant set of skillsets, such as tactics and team play, the idea of PE as a ‘teaching-games-for-understanding’ (TGFU) approach developed as a response to the testing regime in the 1980s (Kirk, 2010). The approach placed more focus on modifying the activity to fit the level of the class and to facilitate learning in other forms than just sporting techniques in football, volleyball, basketball, and so on. While the traditional approach has been preoccupied within a culture of acquisition (of technical skills), the new idea of the idea of PE was more preoccupied with a culture of learning. While the approach of TGFU seems promising from a pedagogic perspective and from a motivational perspective (Hortigüela Alcalá & Hernando Garijo, 2017), Kirk identifies issues for the approach, such as a restricted time-table and the need for PE to structure itself as other subjects to be legitimised in schools (2010, p. 104), teacher competence (p. 47) and a reluctance to change the nature of PE, both due to some pupils enjoying the sporting style of PE and due to teachers in PE having a sporting background which they tend to reproduce as teachers (p. 63). Finally, PE-as-sport-techniques approach, which works in a behaviourist paradigm, is relatively easy to measure and to assess, which according to Kirk has kept PE in this paradigm, while more modern and learning-oriented approaches, such as the TGFU, do not fit within the ‘institutionalised PE’. Boundaries such as the timetable, teachers’ sporting background, demands to documentation and a culture of acquisition makes the implementation of LOA in PE challenging.

PE today is less concerned about discipline and pupil’s readiness for warfare, but rather aims to lay the foundation for lifelong activity for the pupils. The focus has gradually shifted from national defence towards national health (Kirk, 2010; Utdanningsdirektoratet, 2015). The learning plan in PE says that the subject should inspire to a physically active lifestyle and lifelong enjoyment with physical activity. PE is supposed to contribute to a positive view of the body and pupils are supposed to gain knowledge with a wide selection of activities and abilities related to exercise, lifestyle, health and movement in nature (Utdanningsdirektoratet, 2015). However, some research indicates that assessment and feedback in the subject still very much surrounds the context of Kirk’s PE-as-sport-techniques (Leirhaug & Annerstedt, 2016, p. 627), where pupils perform a range of activities and get graded on their abilities in these. Norwegian, as well as Scandinavian PE classes are what can be describes as multi-

activity models (Leirhaug and MacPhail, 2015, p. 628; Annerstedt, 2008, p. 315), which presents a number of varied physical activities for the pupils that are intended to learn about body and health through activity and inspire to lifelong physical activity through teaching a range of activities. This is in line with PE-as-sport-techniques'. Pupils go through a variety of different sports and activities, practice these, before getting tested in their technical and physical abilities in these activities, alongside their effort and ability to showcase fair play.

From research about the historical development of PE in Norway, we can summarise that PE has developed from having a military and a disciplining purpose to be a sport subject to becoming a health-oriented subject today. However, some research suggests that while health and a physically active lifestyle may be the purpose, much of the teaching and thereby the assessment happen in the sphere of PE-as-sport-techniques. While this thesis cannot argue that such a context will hinder the health objectives of the subject, we will see in the section about competence goals that practicing sporting techniques is a small part of what pupils are supposed to learn.

2.2.1.2 The learning plan in PE

The relevant learning plan for PE in this study is KRO1-04 (Utdanningsdirektoratet, 2015), which is the latest edition of the national curriculum that was introduced in 2006, but which now will be replaced a new curriculum in 2020 (Fagfornyelsen). KRO1-04 will still be valid for the school leavers in year 13 up until 2022. While it is interesting to see if there will be changes to assessment in the new curriculum, it is beyond the scope of this thesis to investigate this new curriculum.

The learning plan consists of the guidelines: purpose, main areas, hour distribution, basic skills, competence goals and assessment. The purpose section of the learning plan highlights that PE is a formative subject that promotes a physically active lifestyle and lifelong physical activity. It highlights aspects such as fair play, positive body images, different movement cultures, knowledge about training and nutrition. When it comes to assessment, pupils are supposed to be assessed on their own capabilities and their effort. In the section about main areas, three areas are highlighted for PE pupils in secondary school, (1) sports activities, (2) outdoor recreation and (3) training and lifestyle. Attendees are supposed to have 223 hours of PE in total across three years in secondary school, or a little more than 74 hours annually. The basic skills section states that PE, in similarity with all other subjects, should integrate the core skillsets, oral speech, writing, reading, calculation and digital skills in the subject.

The next section in the learning plan, competence goals, will be explored in an own section of this literature review. The last part of the learning plan in PE deals with assessment.

However, it does not contain much information others than that pupils shall get a grade in the subject. The sparse information about assessment brings us to an important aspect of the Norwegian education system, its decentralised system of teaching and assessment.

2.2.1.3 Norwegian education: a decentralised system

The lack of clear guidelines on assessment in the learning plan represents one feature important for the research context. It makes the Norwegian education system is highly decentralised, where teaching and assessment lie in the hands of the teacher. The decentralisation in education was part of the neoliberalism and NPM ideas in the 1980s, where it is up to local learning plans and the teacher to design teaching and assessment. This came from the MBO idea, where the ‘objectives’ are represented by the competence goals (Møller & Skedsmo, 2013). The decentralised system makes teachers responsible for assigning high-stakes grades that will impact pupils’ opportunities to get into the upper secondary school they want (Leirhaug & Annerstedt, 2016, p. 619). This adds in the importance of using the teachers as reference points for the research, because there are few official guidelines on how to design assessment. Thus, ultimately it is up to each teacher to create assessment designs that promote learning, regardless of what intentions the government department of education may have. There are no standardised tests in PE, so any tests used or not used in the assessment is ultimately the decision of the teacher.

The law related to assessment is the regulations of the education law (Lovdata, n.d.). *Opplæringslova* § 3-2 confirms the pupil’s right to (1) *undervegsvurdering* (formative assessment) (2) *sluttvurdering*, (final graded assessment) and (3) documentation of his/her performance in the subject. § 3-3 states that the assessment should promote learning and give information about the pupil’s competence, both throughout and at the end of the assessment, with reference to the competence goals. § 3-10 deals with formative assessment and states that this can be conducted orally or written. It is built on four criteria of formative assessment, all of which will be discussed in the coming section about AFL in Norway. The regulation about formative assessment represents a strengthening of informal assessment in PE (Engvik, 2010, p. 202), as opposed to more formalised assessment half-yearly and yearly assessments.

Each school, or teacher, develops its own local plan for PE but must work by the competence goals set nationally. In some cases, schools in defined regions work together to create regional learning plans, based on the competence goals (MBO). The Norwegian competence goals represent a criterion-referenced assessment system (Annerstedt & Larsson, 2010, p. 98), where everyone should have the opportunity to get the highest grade if they perform well enough in the competence goals. The opposite is a norm-referenced assessment system, as was the case in Sweden up until 1994, where the grades are scaled and dispersed from high to low in each class. The high-performing pupils in the class would get the highest grades, while the low-performing pupils would get the lowest grade, regardless of their abilities according to the goals. With the criterion-referenced system, the goal is not to rank pupils but to evaluate whether they have reached the intended learning goals (Engvik, 2010, p. 211). As we can interpret from the MBO approach, the competence goals are highly important. However, just as important are the teachers who design teaching and assessment, and who does the interpretation of the competence goals. This may or may not be done according to a locally designed learning plan, but regardless, it is ultimately the teacher who practice and implement it.

2.2.1.4 Competence goals in PE

The competence goals for secondary education in Norway is divided into the three main areas, (1) sports activities, (2) outdoor recreation and (3) training and lifestyle (Utdanningsdirektoratet, 2015) with seven competence goals related to the first one, three to the second and five to the third. Sports activities have more competence goals than the others, which indicate what Kirk (2010) argues, that PE still resides in the PE -as-sport-techniques paradigm. The goals contain both very concrete competence goals such as “*explain and execute lifesaving in water*” but also not so concrete goals such as “*practice and execute various skills in chosen team sports, individual sports and alternative movement cultures*” (Utdanningsdirektoratet, 2015). Such competence goals create substantial leeway for schools and teachers in the selection of activities and what skills to teach and assess. The competence goals also consist of goals that highlights pupil understanding such as “*explaining how different body images and movement cultures influence training, nutrition, lifestyle and health*”. If PE today is characterised by a ‘molecular’ approach that tests sports techniques in pupils (Kirk, 2010), this is certainly not in line with competence goals that highlights understanding and non-physical skills. However, so far in this literature review, we cannot argue that PE today is all about testing, although some literature highlights such

characteristics. While competence goals represent an important aspect of the topic, the main focus in this research is on the assessment practices themselves. To what degree are LOA practices implemented in Norwegian education? The next section will go deeper into this.

2.2.1.5 Report on AFL in Norway

After the introduction of the AFL concept with the Knowledge Promotion, there was a further initiative by the Norwegian Directorate for Education and Training (UDIR) in 2010. The background for the AFL intervention was according to UDIR that regulations on assessment were thought of as unclear, teacher education institutions lacked expertise on assessment, assessment culture was regarded as weak and there was little research on the topic (Utdanningsdirektoratet, 2018, p. 3). The initiative was requested by teachers after the new curriculum came. This resulted in the 'Better Assessment Practices Project' by UDIR in 2007, which looked at clearer regulations and attainment indicators in assessment and which resulted in the Ministry of Education and Research ordering a project set to improve assessment culture and skills and practices (p. 4). The AFL project has also been examined by OECD as a case study *Governing complex education systems* in 2013, about halfway through the project. The OECD title illustrates the reality of such interventions in the education system: that it is complex (Hopfenbeck et al, 2013).

The project aimed to promote AFL in Norwegian education, based on the success of AFL in international studies (Utdanningsdirektoratet, 2010). The goal of the initiative was to facilitate an assessment culture with learning as the main objective, for schools, school leaders and apprenticeship companies. The initiative looks specifically at the key concept of formative assessment (*underveisvurdering*) as a vital means in assessment. The concept is mentioned in relation to final graded assessment (*sluttvurdering*). Key measures in the project includes adjustment of policies and regulations (*Opplæringslova*), public reports about assessment, professional development and knowledge exchange. The AFL project was built on four principles for AFL. Firstly, that pupils should know what is expected of them, secondly, that they get feedback on their work's quality. Thirdly, they should be instructed on how to improve their work, and finally, pupils must be involved in their own assessment. These principles are closely associated with the five strategies suggested by William and Thompson (2007), that are used to guide this thesis, with a few differences. Firstly the four strategies by UDIR lacks the second strategy by William and Thompson about engineering learning tasks that reveal the pupils' current level of knowledge and understanding (2009),

and the fourth and fifth strategies by William and Thompson on peer assessment and self-assessment are combined into one in UDIR's formative strategies (Utdanningsdirektoratet, 2018). As we will see from the results and discussion, implementing also the second strategy might not be a bad idea in Norway.

The program has paid for and organised seminars locally to raise awareness and knowledge levels about AFL. The report states that the directorate did not claim to be experts on the topic and the seminars were a shared process. 73 % of Norwegian municipalities have been represented in the seminars and a total of 1500 schools participated (Utdanningsdirektoratet, 2018, p. 11). UDIR assumes it would take 3-5 years to change education practice, and since participation in the project only lasts for 1.5-2 years, the project has focused on strengthening school capacity to develop AFL practices in the long run (p. 14). In the review of the project, it becomes clear that the knowledge about AFL has raised throughout the project period, particularly when it comes to clarity about the expectations and the learning objectives, where about 50-60 % of the pupils at secondary school are affirmative about this in a questionnaire in 2013 and 2017 (Utdanningsdirektoratet, 2018, p. 17). However, on questions about self-assessment and participation in the assessment process, questions such as "are you able in take part in assessing your schoolwork?" the positive response rate was 17 % in 2013 and 20.4 % in 2017 (p. 18). These are very low numbers and it would probably require much more guidelines, seminars and interventions to elevate these numbers to a point where the fifth formative strategy (self-assessment) by William and Thompson (2009) is a part of the assessment culture in Norwegian schools.

In the decentralised education context, the report describes that it has been challenging to implement practices and to keep participants committed throughout the process and to respond to challenges faced by the participants (Utdanningsdirektoratet, 2018, p. 23). The complexity of the education system and the challenge in changing it from above, through voluntary initiatives, makes UDIR suggest that local school authorities such drive the process towards more LOA practices. The report states that most of the schools in the project have plans to continue with initiatives locally to build stronger AFL cultures. Perseverance and commitment are stressed as important challenges to continued improvement in AFL, as well as holding on to key persons and resources in the schools (p. 25). The availability of online resources for teachers and relevant seminars are also highlighted. An online development course on AFL was developed in 2014 to support the project. The AFL project report states

that there is a danger of the underlying intentions getting lost in the implementation, so-called ‘the spirit vs the letter’ challenges (p. 6). My question to this, is whether the way PE is presented to pupils in Norway mix well with the formative assessment strategies. If molecularised PE, restricted timetables, teachers’ sporting background, documentation needs and other factors limit LOA such as TGFU, what are the implications for AFL? Although the report by UDIR undoubtedly shows the increased awareness about AFL, no one should yet be convinced that assessment in Norwegian PE classes is in line with the formative assessment as defined by Black and William (2009).

In a report about assessment practices in four subjects in Norway, Sandvik and colleagues (2012) interviewed teachers through focus groups and questionnaires, which looked at their understanding of assessment in relation to the learning plan. The report suggests that teachers are more concerned about assessment on a day-to-day basis than about the longer-term learning process (p. 2). Sandvik and colleagues also suggests that school owners are more concerned with reliability than with validity. It is more important to them on a day-to-day basis that the assessment is just and fair, rather than valid. This represent some important questions for this thesis. Are PE teachers happy with creating assessment situations were ability is easy to measure and compare with other pupils? Or do they consider the individual learning of each pupil, by facilitating learning situations where they instruct each pupil? The report also shows that while there are few differences between other subjects, PE is an exception to this. The 31 PE teachers in the study, that teach at a secondary school level, suggests that formative assessment is less apparent in PE than in the other subjects (p. 152) and schools feel more in need of national regulations than in other subjects (p. 144).

2.2.1.6 Conclusion research context

The research context for this thesis is first and foremost a subject that has health concerns as its main purpose, represented by a goal of creating lifelong physical activity. The subject has developed from having a military purpose, to a sports-oriented subject with a focus on teaching multiple sporting activities, to the health-oriented subject it is today. However, although health represents the purpose of PE in Norway, research suggests that the activity and thereby the assessment still very much revolves around the testing of specific sporting skills and more innovative approaches like TGFU are largely hindered by structures surrounding the subject. The learning plan in PE provides no suggestions on what assessment should look like but provides competence goals that represent the MBO system in Norwegian

education. This creates a decentralised system, where teachers have much leeway when it comes to assessment, thereby making the teacher the most interesting perspective for this thesis' research questions. The government has initiated a project about AFL in the Norway in 2010, which has focused on raising the awareness of and competence with AFL practices in the system. However, transforming the assessment culture from assessment of learning, to AFL is complex and while the project has had its successes, reports have shown that there is a lack of formative assessment in PE. The next section will go deeper into the literature about assessment and learning.

2.2.2 Research question 1: assessment practices

This section of the literature review will look at research that highlights the assessment practices used by teachers in PE, mostly in Norway but also in Scandinavia and elsewhere. While the search for formative assessment strategies represent the core of this section, such practices are not necessarily entrenched in PE classes around the world. Formative assessment constitutes a part of what is labelled 'alternative' assessment in the review by Lopez-Pastor and colleagues (2012), as opposed to 'traditional' assessment. In their review, they establish that although there are many 'learning-oriented' approaches and concepts related to assessment in PE, such as authentic, integrated, formative and AFL, such methods are not necessarily the standard in contemporary PE classes. There appears to be structures or factors that hinder the development of more learning-friendly approaches to assessment in the subject. As Lopez-Pastor et al suggest, PE seems to have an inability to adapt (p. 59) and therefore is not aligned with current knowledge about assessment and learning.

In a quantitative study of almost 1500 pupils from upper secondary PE classes in Norway, Leirhaug and Annerstedt (2016) found that the implementation of AFL practices after the 2006 education reform has not been effective (p. 625) and that there were no significant differences between the extent to which AFL-related strategies were implemented across the six schools they researched (p. 625). Likewise, in Australia, Georgakis and Wilson (2012) found in a qualitative study of 17 teachers that AFL was not truly apparent in school PE and sport. The assessment was more concerned with behavioural aspects than with learning and a link between research and practice was not evident. As we will see, these studies are not the only ones highlighting the lack of LOA strategies in the subject. And despite curriculum attempts in Norway (Utdanningsdirektoraret, 2018) to incorporate AFL into Norwegian

education, the decentralised system puts great significance to the interpretation and practice of the teacher.

2.2.2.1 Clarifying and sharing learning intentions and criteria for success

The first formative strategy presented by Black and Paul (2009) is sharing criteria for success with pupils. In Wiken's master thesis (2011) about how PE pupils experience the subject, she found that 3 out of 5 pupils had no familiarity with the competence goals (p. 60). Mørken did quantitative measures in his master thesis (2010) and found in a survey of 82 PE teachers that there is great variance in whether teachers make pupils aware of specific learning goals for each PE session (p. 47). Græsholt found (2011, p. 81) in his master's thesis in upper secondary education in Norway that pupils were not familiar with the learning criteria. In his interviews with ten pupils, with a spectrum of grades in PE, he found that pupils thought it was important that they knew the learning criteria, but that few could actually remember whether the teacher had shared it with them, or not. Later, Leirhaug and Annerstedt's study (2016, p. 622) confirmed that about half of the pupils in a questionnaire of 1486 pupils did not feel informed about the learning intentions from the learning plan, which indicate that this strategy has challenges in Norwegian PE classes. Hansen's thesis (2017) confirms that pupils have limited knowledge about assessment criteria in PE (p. 38). While one cannot expect that pupils remember everything their teacher say, I expect that they at least remember having discussed the criteria. Thorsnæs (2019) did a master's thesis relatively similar to this one, but with focus on the pupils. She also found varying degree to which pupils were informed about or had reflected around the learning goals. Thorsnæs suggested that the presented criteria are not followed up by the teachers, which thereby negatively impacts the perceived usefulness of the criteria for the pupils (p. 46). On the teachers' side of thing, 23 teachers in focus groups thought that these learning intentions were too vague to give valuable information to the pupils (Leirhaug & Annerstedt, 2016, p. 622). The literature seems to confirm that the first formative strategy by William and Thompson (2007) faces great challenges.

2.2.2.2 Engineering learning tasks that elicit evidence of student understanding

The second strategy about the "engineering... of learning tasks that elicit evidence of pupil understanding" (William & Thompson, 2007, p. 8) will be linked to the assessment methods in this literature review. Teachers are responsible for 'engineering' assessment practices, which ultimately decide the degree to which pupils can showcase their learning and abilities in class. For example, if the teacher using PFTs as the main assessment method, the pupil will

hardly be able to showcase fair play. Below, I will briefly look into evidence of testing and authentic assessment in PE, as two broad types of assessment practices, where one is concerned about setting up a testing situation designed to measure ability, while the other is about evaluation of 'real' situations, such as game play. Engvik stated in 2010 that there was little evidence about the use of different assessment tools in Norwegian PE classes (p. 208) and I would argue that the research still is hard to come by. However, there are quite a few references to testing in PE assessment, which gives us a sense that assessment methods might represent a challenge with regards to LOA.

Testing

The use of PFT has according to Lopez-Pastor et al (2012, p. 60) represented a technical view on assessment, with a perspective on PE as a means to train the body. Testing represents an assessment method where pupils' performance or competence is assessed at one point in time. While testing can be ungraded, such as through formative use of summative testing, it typically contributes towards the overall grade in PE, as a way of teachers to test a pupils' skills in technique, fitness or strength. Hence, testing in PE represent a part of high-stakes testing, as the grade in PE at secondary school in Norway contributes towards the overall grade pupils apply to upper secondary school with.

Literature suggest that testing has a role in Norwegian PE classes, although no competence goal in the learning plan calls for it (Utdanningsdirektoratet, 2015). In Græsholt's master thesis (2011), interviews with pupils in PE in upper secondary school in Norway expose that they experience a psychometric testing culture, which focus on what's easily measurable (p. 104). This will be looked at more thoroughly in the section about challenges to LOA.

Græsholt found that pupils thought that assessment in PE could be unfair, that test results were dominant and that boys dominated the classes, to mention a few of the issues (p. 59). Græsholt mentions pupil who felt that the idea of learning in PE and getting graded on this does not happen in today's PE. The participant took up the example of shoot put, which they practice once a year, before being graded in it (p. 61). It is obvious that one session of an exercise does not facilitate a lot of improvement or chance to learn. The same was the case with a cooper test, that pupils went through, without a chance to train for it in advance.

In Leirhaug and Macphail's focus groups with 23 PE teachers, one teacher in the study (2015, p. 630) who was also a math teacher explained great complexities with assessment in PE,

both in practical terms and valuation terms. One difference is the relatively fixed and time efficient assessment of mathematics, where a test is assessed, with right or wrong for each question or task. PE, on the other hand, represent a form of continuous, on-going assessment, mixed with tests or drills and the subjective evaluation of the teacher. There is much flexibility to how assessment is conducted, which creates challenges for each teacher in terms of designing meaningful assessment. In this context, it is easy to understand that testing may represent a simple way of assessing pupils' capabilities, much like a maths test. However, testing may be used in a formative way, for example to aid the teacher and the pupil in checking for progress in a learning process. Leirhaug and MacPhail (p. 630) show teacher stories, where summative testing is being used for formative purposes. The pupils are presented with a drill or a task at the beginning of the course, they are tested on it, before getting tested in the same drill at the end of the course, to give the pupils an opportunity to showcase their progress. However, in the same instance, besides using the skill test formatively, the teacher also explained the motivation behind the use of summative testing for formative purposes was not to facilitate learning but rather documentation purposes. Hansen's master thesis (2017,) suggests, although with only three teacher interviews, there are different opinions when it comes to testing. While one teacher has moved away from it, another teacher has gone to only use it to measure progress (formative use of summative testing), while the last teacher use testing to some degree (p. 43). While much literature mention testing in the context of assessment in PE (Lopez-Pastor et al, 2013; Græsholt, 2011; Hansen, 2017; Jonskås, 2009), there are also other forms of evaluating PE pupils, such as authentic assessment.

Authentic assessment

As mentioned in the clarification of the key concepts in this thesis, authentic assessment represents assessment where pupils conduct realistic tasks (Engvik, 2010, p. 209), rather than decontextualized testing situations. The teacher lets the pupils practice an activity, such as a ballgame or an individual activity and assesses the pupil based on this. In this way, authentic assessment can be linked to observation as an assessment method. Mintah (2003) looked at authentic assessment in US public schools' PE by surveying 396 physical educators. In this context, authentic assessment represents an assessment method that evaluate a series of activities, rather than a single performance, such as a PFT (Lund, 1997, as cited in Mintah, 2003, p. 162). 75 % of the participants indicated that they used some form of authentic

assessment, while the rest were sceptical towards the use of time or did not feel they had enough experience with it. Many used authentic methods such as observation, peer observation, projects and video, while somewhat less used for example portfolio assessment (p. 167). The results also suggested that authentic assessment had positive effects on pupils' self-concepts, motivations and skill development (p. 168), although one can debate the strength of these findings, as they were the educators' perceptions. In the case of using video in authentic assessment, questionnaires and interviews with girls in British dance classes (Cassady et al, 2004) showed that girls were negative towards the use of video in the evaluation, because they were anxious about how they looked and felt disappointed by their own performance when looking at the video footage. This highlights that authentic assessment methods do not necessarily fix the poor learning environment of testing. Chen (2005) conducted research interviews and field observations of PE teachers in the US, where he discovered that only five out of 15 teachers used authentic assessment by engaging the pupils in tasks that elicited evidence of learning. Some teachers used rubrics and check lists to assess the pupils in a performance-oriented climate much in the same way as testing, but through more real situations or tasks. The ten other teachers in the study were according to Chen only concerned with attendance, clothing and participation (p. 175). Chen argues that engaging the pupils in relevant tasks is necessary for the pupils to demonstrate learning and that it is a precursor for feedback that is meaningful for the learning goals (p. 177). Some of the teachers in his study successfully combined teacher feedback and peer assessment with authentic assessment, thereby integrating instruction and assessment, which is fundamental for LOA. However, as mentioned, authentic methods may still create a competition-oriented environment, which as I will discuss later, represent a challenge for learning for many pupils.

MacPhail and Halbert (2010) looked at AFL in Irish education, through a mixed methods projects that looked at case studies where an 'assessment wheel' and 'rich tasks' were introduced as a tool for teachers. The assessment wheel was a sheet that organised evidence of learning and worked as a 'record of learning' for both teachers and pupils. MacPhail and Halbert conducted teacher interviews, focus groups, quantitative surveys and analysis of schools and teacher networks which overall indicated a positive impact for learning. The assessment instrument led to a shared understanding of the role of assessment among teachers and pupils. The rich tasks used in the study also helped contextualise the learning criteria for the pupils. In what Kirk (2010) describes as a molecularised approach to PE, contextualising the learning criteria through rich tasks may represent a major opportunity for assessment and

learning in Norwegian PE. In situations where going through a skill track in football is part of the activity and the assessment methods, it is not surprising that pupils will see their technical skills as the most valued trait, which may cause them to neglect competence goals related to cooperation and fair play. One possible assessment method that can be used to combine assessment of learning and AFL, is what Engvik (2010, p. 209) refer to as ‘performance assessment’. The teacher lets the pupil practice the activity multiple times, whereby the teacher coach and facilitate learning for the pupil, and evaluates the pupil based on the best performance he or she made. In this way, the learning process goes ahead without the interruption of tests, while the teacher assesses through observation.

When it comes to authentic assessment in Norwegian PE classes, Jonskås’ (2009) qualitative interviews with six PE teachers in upper secondary school, presents observation as the main assessment tool used by the teachers, alongside testing. In her interviews, teachers suggest that it is hard to get enough information about the pupils in the assessment, but that tests, measurements, observation and theoretical tests were main methods used. Other methods, such as portfolio assessment in PE, were not used (p. 83) and in general receives little attention in research about PE assessment in Norway. From the three teacher interviews conducted by Hansen (2017), it was found that teachers conducted most of the assessment through observation (p. 38). It is hard to find evidence of how teachers design learning activities in PE for authentic assessment, or to what degree they do this. Are teachers reflecting about this? Do they facilitate situations where everyone is able to showcase their capabilities and their progress in front of the teacher, in an ‘authentic’ scenario? There are many questions related to teachers’ assessment practices in Norwegian PE classes, however, it appears that both testing and observation are common practices.

2.2.2.3 Providing feedback that moves learners forward

When it comes to the third formative strategy about feedback, there are quite a few references that investigates this in Norway. In Mørken’s (2010) master thesis about PE assessment in Norway gave questionnaires to 82 PE teachers. He identifies that only about 4 % of the teachers feel they manage to provide feedback to pupils that guide them toward the competence goals, while 9 % say they manage this to some degree (p. 49). The rest of them were not able to do this. Græsholt’s thesis (2011, p. 106) showed that there are pupils that have little experience with face-to-face feedback from teachers. In Leirhaug and Annerstedt’s paper (2016), most of the 23 teachers in focus groups expressed that they gave pupils

continuous feedback during sessions, while the pupils (60 %) in the same study reported that they had not received feedback on how to improve in the subject more than once, or not at all (p. 624). According to Leirhaug and Annerstedt's the lack of pupil experience with feedback may be a result of pupils not interpreting comments from teachers as feedback, such as technical feedback like "lift your head" or "drop your shoulders" (p. 624). Leirhaug and Annerstedt problematizes feedback as appearing in the context PE-as-sport-techniques, where pupils receive feedback based on technique, movements and actions that will not engage them in rich tasks (MacPhail & Halbert, 2010) or give them an opportunity to reflect upon their own learning and progress (Leirhaug & Annerstedt, 2016, p. 627). This leaves us with the question of whether feedback, which is not recognised by pupils as feedback, still helps pupils towards the learning goals. In Hansen's master thesis (2017), he shows that the teachers felt they gave feedback continuously and that they are better at giving feedback than most other subject teachers (2017, p. 38). This conflicts with what Sandvik and colleagues reported in 2012, where PE seemed to come out worse than other subjects in the implementation of AFL practices (p. 152). Thorsnæs (2019) found in her interviews with Norwegian PE pupils that most of them experienced to get feedback from their teachers and that this aided learning. However, some of them found the feedback vague, which confused the pupils (p. 51). While the results by Thorsnæs seems promising in terms the pupils experiencing useful feedback, there are many questions regarding the frequency and quality of feedback by teachers, that require further research, particularly from quantitative studies.

2.2.2.4 Activating students as instructional resources for one another

The fourth formative assessment strategy is about using pupils as instructional resources for each other, which appears to be of little significance in Norwegian PE assessment. Peer assessment was only conducted by one out of 23 PE teachers in Leirhaug's focus groups (2016, p. 29), suggesting there is a long way to go in the implementation of this. In the same focus groups, but a different paper (Leirhaug & MacPhail 2015), one teacher expressed that his pupils were negative about peer assessment and of putting a grade on each other and that they would not be honest to each other (p. 633). The teacher thinks peer assessment in terms of guiding each other would be much easier. The teacher also remarks that typically, assessment in school is not about guidance, but about numbers and grades, which may be what challenges the use of peer assessment and the implementation of LOA in general. Thorsnæs' master thesis (2019) exposed that pupils had some experience with peer assessment, although not regularly, and that pupils have varying degree of satisfaction with

this. Some pupils confirmed Leirhaug and Macphail's thoughts that they would feel discomfort in assessing peers, while other pupils felt the benefit of getting perspectives from other assessors than the teacher (p. 69). Butler and Hodge (2001) did a qualitative case study looking into what ways peer assessment could improve learning in the US. An open-ended questionnaire suggested that peer assessment could be used effectively when pupils were given instructions and that it could work to create trust and cooperation between classmates. In one of few international studies about peer assessment, Butler and Hodge argue that peer assessment in PE is a great authentic assessment tool that can reduce negative influence of avoidance strategies. They show how pupils that are working in pairs or smaller groups who assess each other, might reduce assessment stress, as opposed to performing an activity in front of the whole class. This research suggests that peer assessment should be effective in Norway as well, while not in the context that the teacher in Leirhaug and Macphail' (2015) study describes, where peer assessment is about numbers and grades, rather than guidance and feedback, which may reduce the positive impact Butler and Hodge describes.

2.2.2.5 Activating students as the owners of their own learning

The last of the formative assessment strategies, related to self-assessment, is much more discussed in research in Norway than peer assessment. Having pupils' conducting more self-reflection in PE is in line with arguments about increasing pupils' self-awareness and self-concepts that can facilitate a physically active lifestyle (Ommundsen, 2013). However, also here we can note some scepticism towards the practical use of it in the literature. Jonskås (2009) found that teachers were hesitant about and did not use pupil self-assessment as part of their assessment (p. 88). Jonskås mentions issues such as the availability of time and the perceived resources teachers have to change their assessment practices, as well as hesitancy towards the ability of pupils to assess themselves. It is necessary to mention that Jonskås' study was performed relatively shortly after the focus on new and more 'formative' assessment strategies in the curriculum reform in 2006. In Græsholt's semi-structured interviews with PE pupils (2011, p. 88), the only experience they had with self-assessment was in the context of evaluating self-organised exercise in PE (*egentrening*), where the teacher looks whether the pupil has assessed him/herself. Græsholt regards this more as a summary by the pupil of his/her own work, rather than a critical self-evaluation, which removes some of the developmental purpose of self-assessment. The pupils only knew about self-evaluation for this activity in PE, while it was not implemented in other activities, which indicates a very narrow use of self-assessment. In Hansen's thesis (2017, p. 44), the three

teachers interviewed let pupils assess themselves, formally through an assessment by one of them, while the others attempted to do this in class. In Hansen's interviews with the pupils however, only two out of six say they have assessed themselves. This tells us that self-assessment is not as entrenched in the system as it should be, from the perspective of LOA. Thorsnæs (2019, p. 58) showed more promising results, where all nine interviewed pupils had experience with self-assessment, while for some of them this was on a rare basis, as also suggested by Græsholt (2011). In addition to this, the pupils experienced some uncertainty regarding the practice, which Thorsnæs suggests may be due to the lack of understanding of the learning goals (2019, p. 61), which highlights the important link between the first formative strategy and peer- and self-assessment.

Græsholt (2011) suggests in his study that self-assessment was more about initiating a 'reality orientation' among pupils, rather than engaging in deeper level engagement with assessment (p. 29). Leirhaug (2016) finds that pupils were not engaged in the assessment process and that understanding of assessment among pupils, so-called 'assessment literacy', was low. Leirhaug and MacPhail (2015, p. 636) found that even teachers that stood out as most innovative when it came to implementing AFL practices, were still in an exploratory phase of trying out these. Measures such as self-assessment were typically introduced late in the course, where pupils were asked about what grade they would assign themselves. If such measures are not used earlier in the learning process to create more active learners, I question the effects this will have on pupil learning, as they only need to be 'owners of their own learning' occasionally. One teacher interestingly explains that there is great correspondence between what grade pupils give themselves and what the teacher gives them (p. 630). The latter point should illustrate a potential for using self-assessment more, as pupils seem to be honest when they assess themselves. The evidence from the Norway suggests that self-assessment, despite the AFL objectives of the government, is not deeply entrenched in assessment in PE and even when it is conducted, it is more a matter of grading themselves rather than engaging the pupils in deeper reflections about their own performance and learning. This tells us that although there is increased focus on AFL practices in Norwegian PE, some of these practices remain shallow.

2.2.2.6 Conclusion assessment practices

Although much of the literature do not look at assessment practices specifically, much of it still reveals something about the assessment practices in Norwegian PE. When it comes to the

first formative strategy by William and Thompson (2007), there seems to be variance in what degree pupils are informed about the goals and expectations of the subject. While some research suggests teachers inform about this, pupils do not always agree. Regarding the second formative strategy, assessment practices seem to revolve around testing and authentic assessment, such as observation. Much of the assessment still seem to locate in the PE-as-sport-techniques paradigm, or at least have a technical perspective on skills in the subject. The third formative strategy, about feedback, also seems to have different findings in the literature. While some evidence show that teachers do not feel they have the time to give regular feedback to all pupils, there is at least a growing amount of literature that discusses it and finds evidence of practice. Peer assessment, the fourth strategy, receives little attention in literature, while the fifth strategy, self-assessment, is used, but does not seem to engage pupils in deep levels of self-reflection.

2.2.3 Research question 2: challenges to LOA

This section of the literature review looks at challenges to assessment and the implementation of LOA in PE. The main search area is Norway, but also international research is included. Very little research discusses the challenges to implement formative assessment, and in Norway, the search for such literature is quite challenging. While the Norwegian government has had ambitions about the introduction of AFL since 2006, changing education systems takes time (Berry 2011) and I assume there are structures that make the shift to more learning-friendly practices challenging. The review by Lopez-Pastor and colleagues (2012) brought up the need to explore the barriers to implementing LOA in their suggestions to future research (p.71). This suggestion motivates this thesis, as it strives to understand the challenges with assessment in Norway, seen from the teacher's perspective. Leirhaug and MacPhail (2015) suggests that for the PE teachers, the surrounding structures of the subject caused innovative teachers to 'explore, rather than implement' AFL practices (p. 637). Leirhaug and MacPhail are encountered with the tension between formative and summative assessment in their study, where teachers might be positive towards implementing formative assessment practices, but where such practices are compromised by documentation needs and summative structures. This shows evidence of Kirk's suggestion (2010, p. 104), that PE struggles to confine itself within school structures, where the timetable and evaluative requirements ultimately conflicts with its purpose of promoting lifelong physical education. Leirhaug and MacPhail are also concerned about teachers' assessment practices being framed in PE-as-sport-techniques. They suggest that "the potential of AFL ... vanishes, replaced by

what looks like assessment for assessment's sake" (p. 637). They call for a shift of focus to the individual learner and learning experiences, thorough planning and support for pupils as responsible for their own learning. This calls into question the view pupils have on PE.

2.2.3.1 Pupils' view on PE

While assessment can have negative effects on pupils' learning, another challenge is the view that the pupils themselves have on learning. It might be a challenge when it comes to the implementation of LOA, if pupils do not plan to learn anything. In Denmark, Seelen (2012) discovered in his doctoral thesis that pupils' attitude towards learning in PE was non-existent. They did not perceive PE as a learning subject and did not think they learnt anything in the subject. The pupils saw PE as a subject where being active and having fun was the main purpose (p. 178). In Norway, Wessel's master thesis (2014) describes the anticipating factors by pupils that teachers face when they design their teaching and assessment. Wessel describes (p. 82) that anticipation by pupils, where they expect 'fun' games like bandy, basketball and football. The same was shown in Kolbeinsen's master thesis (2016), where pupils wanted more "normal" PE, meaning playing games and sports (p. 45). Similarly, in Hansen's master thesis, pupils saw PE as an activity subject (2017, p. 36), as opposed to a learning subject. Wiken (2011) found that 60 % of the participating pupils in the last year of upper secondary education PE (year 13) thought that they had learnt nothing in the subject (p. 61). As she notes in her discussion, this is disturbing considering that these pupils have been through 13 years of PE. On the contrary, many participants in the study experienced the subject as a playing ground without specific learning goals and 1/3 of them felt PE was less important than other subjects (p. 70).

If pupils do not see PE as a learning subject, this raises some critical questions regarding the implementation of LOA strategies. How do pupils' attitudes influence the teacher's teaching and assessment practices, as well as the types of activities they do? Andrews and Johansen (2005) showed that often the pupils often practically decide the activity themselves in PE, whereby boys dominated the discussion and got their will when it came to the choice of activity. This may suggest that when pupils have influence on how PE classes are designed, this may have negative effects on the inclusiveness of the subject, which affects the learning environment. There are also questions such as: are pupils responsive to feedback designed to forward their learning, or will they neglect it? How easy is it to engage pupils with meaningful self- or peer assessment if they see PE as merely a subject for playing sports?

2.2.3.2 Competitive environment and its effect on pupils

The lack of pupils' interpretation of PE as a learning subject may be due to the emphasis on sports activity, through the PE as sport-techniques, as Kirk suggests (2010). Annerstedt's comparative textual analysis (2008) revealed that the Scandinavian countries are very similar in PE and highlights to emphasis on the multi-activity model, where pupils are supposed to learn a range of activities. Annerstedt confirms what Kirk suggests for contemporary PE, that certain activities dominate the field, such as ball games, athletics, gymnastics and fitness training (p. 313), as well as winter sports in Scandinavia (p. 315). In Wiken's survey (2011) of 228 upper secondary PE pupils, these saw physical skills as the most valuable traits for assessment and 2/3 of them felt like being tested physically in the subject (p. 64). The activity context of PE and the way it facilitates participation for pupils have impacts for assessment and learning. One of the challenges to implementing LOA in Norwegian PE classes is the diversity of pupils in PE (Hansen, 2017, p. 35), which creates challenging implications for the implementation of assessment that facilitates learning for whole classes.

Three different groups of PE pupils have been identified through a survey by Ingebrigtsen and Mehus (2006) of 1816 Norwegians in secondary school. These are firstly, competition-oriented pupils, secondly, activity-oriented pupils and finally, the less active pupils. The categorisation of pupils suggests that different pupils have different needs and desires for the PE subject, which represent a challenge for teachers, as different groups of pupils may respond differently to the same assessment strategy. If the teacher uses testing in the assessment, this will undoubtedly favour the competitive pupils and those that excel in sports (Säfvenbom et al, 2014; Leirhaug & MacPhail, 2015, p. 635), while it may cause less active pupils to engage with avoidance strategies or make them dislike the subject (Andrews & Johansen, 2005). Kolbeinsen (2016) mentions a pupil that stopped attending organised sports because of the competition-environment and who felt the same environment again in PE classes. The pupil said the teacher had a background from competitive sports and took the role of a 'sports coach' in PE sessions, where many pupils felt a performance pressure (p. 49). In addition to this, pupils in PE spend much time comparing themselves to others in the class (p. 52) which creates challenges for the teacher to design a learning environment, rather than a competition environment. Activities that make it visible who is a high-performer and who is not, is an example of practice in PE that may make it difficult for many pupils. The same goes with activities where it is easy for a few pupils to dominate (Seelen, 2012, p. 200).

Andrews and Johansen (2005) conducted interviews with 13 girls in upper secondary school based on a questionnaire, where they had expressed their dislikes for PE. The study found that the girls experienced discomfort with being graded and 'monitored' by the teacher (p. 304). They experienced a performance pressure when they reached secondary school. They also confirm that activities are mostly chosen on the boys' premises and they play football more than other activities (p. 306). The study reveals that the girls used avoidance strategies in PE activities because they were afraid to perform bad in the class, thereby risking the loss of popularity socially (p. 310). There are also locker room challenges such as avoidance to showering in front of others (p. 308), which have negative impacts for participation and thus create challenges for assessment. Andrews and Johansen's survey were conducted 18 years ago, however, in interviews and focus groups with six girls from upper secondary PE, Omholt (2015) confirms that girls think that PE is more designed for boys in current practice. They highlighted the lack of non-traditional sporting activities such as dance, various forms of play and outdoor recreation. The lack of prioritisation of such activities was also highlighted by Jonskås (2009). Team and ball sports dominate the classes according to the participants in her interviews. Lagestad's master's thesis (2017) investigated grades of upper secondary PE pupils in Norway, where boys received significantly better grades in PE than girls. This was due to boys getting much more top grades (grades 5 and 6) than girls and suggests that PE is designed in a way that better suits boys than girls, on a system level.

The issues with groups that struggle with PE, or find that certain groups dominate the subject, represents a challenge for LOA in PE. There seems to be a competition environment out of line with learning-oriented strategies such as TGFU. In addition to this, effects such as avoidance strategies means that meaningful assessment of many pupils will be very challenging or impossible. There appears to be great challenges to the second formative strategy by William and Thompson, which was not among the formative strategies implemented in the Norwegian AFL project (Utdanningsdirektoratet, 2018). Formative assessment strategies may, however, represent opportunities to change this, but not without changing the nature of the classes, such as the competition-oriented activities that cause some pupils to undertake avoidance strategies. Such strategies are closely linked to pupils' motivation in PE.

2.2.3.3 Motivational implications for pupils

The competitive environment that seems to entrench PE, requires some discussion around the motivational effects on pupils. Seelen (2012) highlights that the lack of formal assessment in the subject, such as examinations and national tests, provides the teachers with much power in designing the assessment. Implementing such formal assessment is not ideal from an ideological perspective, as poor results on examinations or tests may negatively impact pupils' sense of competence and thereby their motivation to be physically active (Ryan & Deci, 2020; Seelen, 2012, p. 181). For the pupils that are in the less active categorisation of PE pupils (Ingebrigtsen & Mehus, 2006), one question is how to motivate these pupils. If the subject is going to reach its goals of facilitating lifelong physical activity for its participants, creating a motivating environment is crucial. In addition to this, it does not matter if LOA is achieved, if pupils do not participate. As Harlen states (2012, p. 172) it is not necessarily obvious for the pupil that making an effort in PE create positive value for him or her.

The grade in PE will impact the pupil's opportunity to get into the upper secondary school he or she wants. This creates a good possibility for extrinsic motivation (Ryan & Deci, 2020) to be the dominant type of motivation in the subject and in school in general. In this way, pupils may steer towards performance-orientation, rather than learning-orientation, which challenges LOA, as it induces passive learning strategies and avoidance of challenges (Harlen, 2012, p. 176). This collides with the second formative strategy, about creating effective classroom discussion and it will also challenge the engagement of pupils as owners of their own learning. Pupils expect an external assessment from the teacher, where the extrinsic goal of performing goes at the expense of intrinsic goals for the subject. Teachers may also harm the motivation of pupils, by providing feedback that is judgmental or grade-driven, may have negative effects on the self-efficacy of the pupil and the motivation to learn (Harlen, 2012, p. 177). Harlen brings forward two rules for assessment that must be in place to preserve a learning culture (p. 178). This is firstly, assuring that tests are consistent with the pupils' capabilities and secondly, to promote self-comparison rather than comparison between pupils. The first rule is important for avoiding pupils' loss of self-esteem and self-efficacy, which can result in avoidance strategies (Andrews & Johansen, 2005). The second rule is important to avoid a competitive environment, which research suggest is present in Norwegian PE classes (Wiken, 2011; Kolbeinsen, 2016). Ommundsen and Kval (2007) investigated motivation and autonomy-support among PE tenth graders in Norway by having 194 participants conduct questionnaires. They argue that performance climates may increase

the prevalence of amotivation among pupils, and that this may be counteracted by emphasising autonomy and mastery-support to increase perceived competence among pupils (p. 405). The positive results of emphasising autonomy, and satiating pupils' needs according to the SDT, has also been confirmed in another Norwegian master's thesis about PE pupils (Nesheim, 2008). The PE-as-sport-techniques context and the disadvantages it creates for some groups of pupils require action, in order to implement LOA. The upper secondary pupils in Kolbeinsen's master thesis (2016, p. 46) said that they had not learnt new movement activities in PE, but that they were interested in doing this and they suggested that it may make the participants more equal and reducing the dominance of some groups in PE (p. 47). 'New' activities make pupils more equal and avoid dominance of 'good' sports pupils, thereby creating learning opportunities for all. This facilitates formative assessment strategies such as creating meaningful subject discussions, peer assessment and engaging pupils as instructional resources for one another. Literature suggests there is a long way to go before accomplishing this.

2.2.3.4 Documentation needs in assessment

Another potential challenge to LOA, mentioned briefly in the literature, is the documentation aspect of assessment in PE. The focus on documentation in PE, as required by § 3-2 and § 3-35 in the *Opplæringslova* regulations, seems to have some negative impacts of the role of teachers in assessment. In research interviews with PE teachers, Arnesen and colleagues (2013, p. 24) exposed that documentation needs can lead to the teacher taking a background position where he can observe, take notes and document performances, rather than taking an active role where he can give feedback and coach pupils. Leirhaug and Macphail (2015, p. 630) presents the story of a teacher who is experimenting with the formative use of summative practices, through a skill track, which he presents once to his pupils, then once again at the end of the school year to check for improvement. While this suggests positive evidence for a focus on learning and AFL, the teacher explains that the main motivation behind this was not AFL, but the need to document pupils' performances. The teachers in Jonskås' study (2009) thought it was hard to collect evidence for providing grades in PE and tended towards using traditional methods of assessment, which could more easily evaluate measurable performances (p. 81). This included testing, observations and theoretical tests. The methods described by the teachers were more in line with assessment of learning, than assessment for learning. As evidence suggests, the need to document pupils' performances as required by national regulations, may lead to the teacher taking a background position

rather than taking part and instructing pupils with feedback throughout the session. This role as an ‘external assessor’ may come in conflict with the teacher’s role as a learning agent, which challenges LOA. However, literature on this is sparse and more research is required.

2.2.3.5 Interpretation of guidelines and resistance to change

The wide perspectives of the learning plan in PE represent a challenge for teachers, as there are many learning goals to consider (Engvik, 2010, p. 203). The goals of PE are both linked to formation objectives, such as cooperation and fair play, and to health objectives, such as lifelong physical activity. In assessment, the teacher firstly needs to consider technical factors, such as the available methods for assessment. Secondly, ideological factors, such as the purposes of PE and thirdly, theoretical factors such as the learning plan and theory on assessment (Eggen, 2008, as cited in Engvik, 2010, p. 203). This can be illustrated by the following example: a teacher that has a technical focus on PE uses testing to evaluate pupils’ competence. He has no trouble assigning grades but is unable to facilitate learning for his pupils. A second teacher has an ideological perspective and focuses on promoting lifelong physical activity. The teacher struggles with assigning valid grades to pupils but do not worry about this, provided his pupils experience enjoyment and mastery in PE. A third teacher has a theoretical perspective and focuses on the competence goals. This teacher has a hard time finding a balance between assessment of learning and as AFL. The decentralised system that provides teachers with a lot of freedom, also creates space for assessment practices that are out of line with formative strategies.

In Australia, Hay and Macdonald (2010) showed in semi-structured interviews and observations in two different schools that the teacher had a huge impact on the construction of ‘ability’, despite of the official guidelines from the syllabus (p. 9). They found in their interviews’ indications that pupils’ habits and possession of certain valued capital, such as the participation in organised sports outside school hours. Such characteristics were beneficial for the pupil regardless of their performance in class (p. 16). This tells that despite their being official guidelines that regulated assessment, the PE teachers still assessed by their own criteria, which speaks of an inability of PE to reform its assessment practices (Lopez-Pastor et al, 2012, p. 59). While it is obvious the teacher’s subjective assessment and discretion represent a fundamental part of PE assessment (Engvik, 2010, p. 212; Birch, 2016), it also creates some challenges for the implementation of formative assessment strategies. If the teacher at an early stage ‘see’ what grade the pupil deserves, what implications does this have

for the interpretation of learning in PE? If the teacher define athleticism as the overarching skill in PE, what will this mean for formative strategies such as feedback? If the teacher has too much personal influence over the construction of ability, regardless of what the competence goals say, this may represent a challenge for what the teacher view as learning in the subject.

There is great variance in how teachers understand assessment regulations in Norway, which creates uncertainty regarding the implementation of LOA practices. In focus groups with 23 PE teachers in upper secondary school, Leirhaug and colleagues (2016, p. 27) discovered that there is great variance in the way teachers understood the competence goals, with implications for how they designed local learning plans. Mørken's thesis revealed that more than 45 % felt they should know more about the goals and 22 % thought the goals were unclear (p. 51). 57 % did not think the instructions provided a good basis for assessment (Mørken, 2010, p. 52). These findings suggest that PE teacher not necessarily know the regulations about assessment that well, many are unsure about these and most teachers are not happy with them. Jonskås (2009) conducted qualitative research interviews with six PE teachers in upper secondary school in Norway. The more experienced teachers thought the guidelines were too vague, while the newer teachers had nothing to compare it with. The older teachers found the 2006 curriculum ambiguous and there were notable differences to what degree curriculum reform had effects on their assessment practices (p. 71). The sample of six teachers is however quite limited to add too much weight into this, however, other have suggested the same (Arnesen et al, 2013). All three teachers in Hansen's thesis (2017) wanted more concrete criteria to the learning goals in the subject (2017, p. 46). The same appears to be the case in Sweden, where teachers in a study by Annerstedt and Larsson (2010) have complained about the lack of clear and concise directions from the government and a lack of training when it comes to assessment, resulting in uncertainty about what they should value in their assessment (p. 107). Arnesen and colleagues claims that the curriculum changes PE in Norway has gone through from 2006 till 2012 have ambiguous effects on assessment (p. 24), despite its increased focus on AFL (Utdanningsdirektoratet, 2018). For some teachers, clarification of the learning objectives has been helpful in the assessment, while for others, it has led to a testing-oriented assessment regime (Arnesen et al, 2013). Leirhaug and Annerstedt (2016) found that teachers in their study were aware of changes in the national curriculum and they adjusted their practices according to their interpretation of the curriculum (p. 626). This provides some optimism for the potential of improving assessment

practices from a top-down approach. However, Leirhaug and Annerstedt claims that the teachers may not have a deeper understanding of AFL practices and that fully implementing LOA to their classes will require much work and changing of current practices. Things brings us to reflect upon the role of teacher education as a potential challenge to LOA.

2.2.3.6 Teacher education: how well prepared are teachers?

If we look beyond assessment guidelines, the training and education of PE teachers may also represent a challenge for LOA. Leirhaug and Annerstedt (2016) concludes that in the introduction of AFL in Norway, teachers did not have the necessary understanding of AFL and the implications it would have on pedagogy and how they organised their PE classes (p. 628). In Mørken's questionnaires (2010) to 82 PE teachers, more than 55 % tried to give feedback but were unsure whether they were able to, while 12 % said it was too hard to give feedback (p. 49). Engvik (2010) notes that formative assessment in PE is challenging for teachers because it puts substantial pressure on their professional competence. PE is a diverse subject that deals with many sports, movement cultures and activities, all of which are distinct fields themselves. It requires huge amounts of knowledge and competence for PE teachers to have the abilities to provide meaningful feedback in all activities that PE constitutes. In this context, it is not surprising that PE teachers may reproduce their own sporting background in their teaching (Kirk, 2010, p. 63; Moen, 2011, p. 234). Moen noted in her dissertation about PETE (2011) that all the students in the research had a shared love of sports. However, this was not necessarily synonymous with competitive sports, as the students admired both competitive and non-competitive aspects of sports (p. 180). Moen also shows that there is a view within PETE that sports skills is the most important skill to develop for students during PETE (p. 230). The pupils that participated in Wiken's questionnaire of upper secondary schools (2011) had less knowledge about areas such as outdoor recreation, dance and orienteering, which were areas that seemed to be less prioritised. Wiken discusses whether the lack of prioritisation of these activities can be related to lacking teacher competence in these areas (p. 61). This evidence suggests that PETE in Norway might reproduce Kirk's (2010) idea about PE-as-sport-techniques in the education of future PE teachers. More research is needed on how and how much contemporary PETE focuses on AFL. The focus on AFL is still quite new in Norway and hopefully PETE today focus more on this.

2.2.3.7 Class sizes and time to assess

When it comes to more physical challenges to implementing LOA practices in PE, few hours and many pupils are also brought up by Engvik (2010, p. 202) as a challenge for LOA. Mørken (2010) found in his questionnaires of 82 PE teachers that almost 20 % say they do not have time to give feedback (p. 49). One teacher interviewed by Jonskås (2009) described that pupils get feedback, but not for all activities. Another teacher described the difficulty in providing continuous feedback to everyone, when he was responsible for ten classes and 300 pupils (p. 55). When Hansen (2017, p. 48) interviewed three teachers concretely about challenges in assessment, one of the biggest challenges brought forward was the class sizes and the difficulty in providing meaningful assessment to all pupils. There is not enough time to follow up on everyone in the class, to the degree that formative assessment strategies require. This represents a major time challenge for the teachers. Also, the decentralised system and the relatively vague competence goals in PE necessitates a ‘community of interpretation’ (Engvik, 2010, p. 205), where teachers and professionals can meet and create standards linked to the competence goals. Hansen’s master thesis discovered however, that the lack of time and the pressure from other subjects hindered seminars and meetings between PE teachers at the schools (2017, p. 36), with negative implications for the community of interpretation that Engvik describes.

Macphail and Halbert’s study in Ireland (2010) looked at the implementation of AFL practices in PE. The key concern for teachers in the study was the time resources needed for planning and implementing the assessment instruments in their study. While they thought that classes worked better with AFL frameworks, there was a considerable time cost associated with this both in classes and in the preparation (p. 36). The time needed to both plan and implement LOA strategies may be one of the biggest barriers to implementing such practices in PE. Traditional approaches, such as physical testing or skill tracks, are very effective and ‘industrial’ approaches that facilitate mass-evaluation of large groups of pupils in PE classes. Focusing more on AFL will for example necessitate more time to coach each pupil in line with strategy number two (William & Thompson, 2007) and provide more feedback that propels learning in line with strategy number three. Turner and colleagues (2017) sent a survey to American schools about resources for the PE subject, which looked at PE teacher credentials, teaching load and budgets. The study revealed a median pupil to teacher ratio of 400 pupils per full-time teacher, which gives an impression of the potential teaching load in the subject, where assessment do not seem to be straightforward. While it is beyond the scope

of this thesis to investigate the teaching load of the participating teachers, the evidence so far suggests that time to assess may be a barrier to the development of assessment practices in PE. The high pupil to teacher ratio and the evidence on PFT practice in Turner and colleagues' study, demonstrates how PE assessment can be linked to neoliberal influences in school (Macdonald, 2011; Møller & Skedsmo, 2013), where assessment is conducted in an industrial fashion. Bezeau and colleagues (2020, p. 389) confirm through interviews with Canadian PE and health teachers that the time resources and the amount of paperwork limit the teachers in assessing the pupils' competence. In their study, teachers were supposed assess pupils' ability to adopt a "healthy and active lifestyle" (p. 390). This is difficult considering that lifestyle is something that happen outside of school and without a good framework, it will be very hard to assess in a school-context. The issue is relevant in the light of Norwegian PE's intended purpose around lifelong physical activity, where wholistic skillsets and deeper understandings of the subject's topic are required, but where the time, resources and methodology seem to limit the teachers' ability to assess these aspects.

2.2.3.8 Conclusion challenges to assessment

This literature has shown that pupils do not always consider PE as a learning subject, but rather expect to merely play sport. The PE-as-sport-techniques context and competition-oriented assessment practices represent challenges for some groups of pupils, where amotivation and avoidance strategies represent a real challenge to LOA. In addition to this, documentation needs seem to take time away from instruction and feedback for teachers. Much research deal with varying teacher practice and teachers' interpretation of official guidelines and learning objectives. It becomes clear that this represent a challenge for reforming the assessment practices in the subject and implementing LOA practices, particularly since there are few guidelines that steer teachers in this direction. Research also suggest that PETE has worked in favour PE-as-sport-techniques, rather than more learning-oriented strategies, however, literature on this is limited. Finally, class sizes and the time PE teachers have for assessment seem to be major challenges for the implementation of LOA. However, it also becomes clear that the literature about challenges to LOA, is quite limited. While some literature discusses this briefly, no literature I have encountered has looked at this specifically.

3 Methodology

3.1 Research design

The choice of methods for a research project should reflect the goals of the project. This thesis aims to illuminate different teacher practices and the challenges that teachers face in assessment. Thus, qualitative research, which is more concerned with “words rather than numbers” (Bryman, 2016, p. 375) qualifies as an appropriate approach. To investigate teachers’ assessment practices in PE and the challenges in LOA for the subject, this research takes an inductive approach and investigates qualitative data that can illuminate the research questions. As this project seeks to expose examples of assessment methods, as well as revealing challenges to LOA in Norwegian PE classes, the natural point of reference will be the teacher. In this way, this research can be described as interpretivist, where we try to understand the world through the perspective of the research participants (Bryman, 2016, p. 375). In order to illuminate my research questions, I need to interview teachers with responsibility for teaching and assessment in Norwegian secondary education schools. To fully understand the complexities of real-life assessment and the everyday challenges that may hinder LOA, the project depends on this perspective. To illuminate the assessment practices and the challenges related to LOA in PE in Norway, qualitative interviews have been conducted in eastern Norway with a wide selection of teachers from a wide selection of schools. Qualitative research interviewing is the best approach to investigate my research questions, as teachers can speak freely about the topic to me as a researcher, with some pre-set questions, but without the rigidity and non-malleableness of a quantitative approach. Eastern Norway was chosen as the research area based on both great variances when it comes to school and city sizes, as well as ease of travel between sites.

3.2 The sampling process

The sampling process in this research represent a form of ‘purposive sampling’ (Bryman, 2016, p. 408). Where ‘probability sampling’ represent a sampling process where randomised samples are selected, purposive sampling speaks of a process that is strategic about getting samples that fits with the research questions. There have been two levels of sampling (Bryman, 2016, p. 410) in the research, one for the schools selected and one for the teachers. Regarding the first level by Bryman, it has been the goal to sample a diverse selection of schools in Eastern Norway. Diverse in this sense means schools in larger cities, smaller cities and in more rural areas, but also diverse in terms of including public schools and private

schools, both with an academic focus and with a sports focus. Obviously, the public schools represent the largest group of schools included in this study, as public schools by far outnumber private schools in Norway. It was a conscious part of the sampling process to include private schools, both in order to include a great variance of schools in addition to getting some impression of the differences between public and private schools, if there are any. For the second level of sampling, it has been a goal to include a diverse group of teachers, to avoid a homogeneous group of interviewees with the risk of having too similar perspectives. The interviews have an equal amount of male and female teachers, and a balanced selection of younger and older teachers. Bryman refers to the process as ‘generic purposive sampling’, which is not motivated by theory, but by a quest for certain types of participants, for example from both young and old teachers. While there is no underlying theory that suggests the importance of this diversity, I imagine that different contexts and different teachers face different challenges in their assessment. In order to maximise a positive response rate from invited participants, invitations were in five cases conducted by meeting up at the schools to see the school leaders or receptionists to request PE teachers. There is obviously a potential for not being sent to certain PE teachers, that the school leader/receptionist for some reason wants to keep away from such a study. However, it is not my impression that this has represented a real problem for the sampling process, as anonymity has been ensured both for individuals and for schools. In addition to that, I have had respondents that have spoken freely about negative impact of school leaders, which leads me to the conclusion that this has not been an issue for the sampling.

The sample size for this research is twelve teachers from ten schools. While it is impossible to know the optimal sample size for this particular research context, this thesis follows the findings of Guest et al (2006), who found that there was a ‘thematic exhaustion’ and a ‘data saturation’ around the twelfth participant. This means that after twelve interviews, the findings in their data would not change substantially after analysing further interviews. They had reached a point where the variety in their sample was reflected in their findings. The research sample in their research was homogeneous and the context very different from this research, so it is impossible to know whether this is a correct sample size for my research topic. However, although anecdotal, I realised there were fewer and fewer surprises in the research interviews for this research project. In the final interviews, I noted that answers to my interview questions were relatively similar to what I heard from other teachers.

3.3 Research interviews

3.3.1 Qualitative semi-structured interviews

This research is guided by qualitative semi-structured interviews that seeks to understand the research interviewees, rather than to stick to a rigorous and strict interview guide that are more concerned with reliability between interviews (Bryman, 2016, p. 466). The interview guide has been flexible and was changed and printed out in new versions four times in total. The reason then for changing the interview guides during the process, despite having small differences is the questions asked during the interviews, is that has increased the opportunity for more relevant and interesting responses seen in the light of the research questions. This was a planned strategy which has allowed me to adjust and improve the research guide according to issues and thoughts that has come up during interviews with teachers. The adjustments have come as a result of what teachers have said or brought up during interviews, but also as a result of questions I have come up with during interviews, such as the number of activities pupils are assessed in throughout the PE subject, which give some sense of how deep or superficial the learning process has been.

The danger of this is that there may be aspect of the research questions that have not come up during the first few interviews, as appropriate questions were introduced later in the process. However, the research purpose of this project is more a matter of understanding the assessment and the challenges related to it, rather than strict comparison between the participants. The overall aim of the research is to bring forward the assessment practices used by PE teachers and the challenges they meet in assessment. This is more important than keeping the exact same interview guide between different interviews. It has been more important to adjust interview guides to make them more relevant, than to stick to the initial guide. However, as Bryman suggests (2016, p. 469), there needs to be some structure to the interviews in a multiple-case setting, to ensure comparability between cases. This was ensured through the semi-structured interview guides, where most of the questions and all themes remained throughout the different interview guides.

One key aspect of the interview guide was to avoid having too specific questions (Bryman, 2016, p. 471), which may have left out interesting perspectives and responses that illuminate the research questions. In this regard, too specific research questions may also become too inflexible, that would leave the research vulnerable if new important perspectives would arise during the interview process. The interview questions also needed to be easy to understand, to

avoid misunderstandings. It was of great importance not asking leading questions, or make the interviewee feel obliged in any way to respond in an affirmative way. The job of the researcher is to ask questions that facilitate open-ended and honest answers and avoid pushing the interviewee into a corner. In hindsight, many of the questions in the interview guides are strictly not important to answer the research questions and may have been left out to make to interviews even more focused. On the other hand, these questions have helped me as a researcher to understand more of the research context. One question was about what the teachers' value in PE and what they understand as learning in the subject. While these parameters at first sight appear not to be relevant to the research question, they nevertheless provided me as a research with knowledge on the topic. As Lopez-Pastor and colleagues suggested in their review (2012, p.72), a strong valuation of certain factors such as fitness or pupils' learning will result in these aspects being assessed more than other aspects. The teacher's personal belief is likely to matter regardless of what the learning plan says.

3.3.2 Interviews

Before the real data collection started I conducted a trial interview with a friend who works as a PE teacher, who otherwise would be a useful participant in the data collection, but whom I left out of the study in case the interview setting with a friend would be conducted differently than the other interviews. Nevertheless, this trial interview provided me with experience before the rest of the interviews and gave me confidence in my research guide, as well as opportunity for adjusting it. A test interview gave the opportunity to get feedback from the interviewee to correct questions that may be irrelevant or insufficient to elaborate on the research questions, as well as to give the opportunity to add questions that may be of interest.

The research interviews took between 45 minutes and 1 hour, while the longest interview was about 1 hour and 20 minutes. The different range in duration can be attributed to the talkativeness of each teacher, as some answered relatively short and concise, while others elaborated much more on each topic with examples and explanations. While the first group was easier in terms of transcription and analysis, the latter group of teachers gave much needed insights to the inner workings of assessment, which is what this research is after. All interviews except one was conducted on the respective schools, in meeting rooms, offices and empty classrooms, while the last one was conducted in a study room at a university library, due to limited space in the interviewee's school. The noise level was kept to a minimum, both to hear and understand all aspects of the interviewees' responses, as well as to make the

recordings loud and clear, to avoid confusion during the transcription. Doing the interviews in the schools have also allowed research participants to participate in an environment that is well-known to them and it made it easier to recruit participants. The research interviews were recorded, both to catch everything said and to be mentally present during the interviews. It was vital to be alert and to be able to follow-up on particularly interesting aspects of the interviewee's responses, and this alertness may be compromised if I had to take notes on paper during the interview (Bryman, 2016, p. 479). On the negative side of things, Bryman notes that (p. 480) the presence of a recording device may make the interviewees more reluctant to speaking freely in the fear of consequences, despite assurances about anonymity. While it is impossible to know whether some teachers have held back information, it has been my impression that most teachers have spoken freely and have not seemed worried to speak out in front of a recording device.

3.4 Transcription

In order to make the research data easier to analyse it was necessary to transcribe the interview recordings to written documents. The transcription has been conducted manually with the support of 'f4transkript', available from UiO, which has given some valuable technical functions in the transcription process, such as adjusting the speed of the recording, pausing, moving back and forth in the recording and typing and listening in the same software. Nevertheless, the transcription was done manually, which has given the opportunity to engage more with the research data, as well as to take informed decisions about how to transcribe and translate data to preserve the original meaning of the raw data. I have translated phrases and Norwegian dialects and sociolects into a common language for all the interviews to ease the analysis process. In doing this, there is a danger of translating out important meanings or words from the responses. Particular concepts and dialectical words can mean different things in different places. Still, as a researcher I feel I have a good overview of the Norwegian language, and the data was collected in a part of Norway where people can understand each other very well.

The downside of the transcription process is without doubt the time consumption. Each interview took four to five hours to transcribe. The alternative to transcribing manually was automatic transcription programs, or an external transcriber, however, these represent some issues both in terms of reliability, validity and financial cost. Also, by using automatic transcription programs, I would as a researcher be deprived on an opportunity to dig deeper

into the research data. In addition to this, necessary translation could have needed to be done by me as a researcher anyway and the software may have misinterpreted words here and there. Finally, in terms of data management and privacy, doing the transcription myself results in one less person being involved with the personal data of the research participants. Another method necessary for this research is the translation of data and findings from Norwegian to English for presentation in the thesis. All research interviews have been conducted in Norwegian, to ensure that practices and issues are explained as accurate as possible. The thesis however is English, which necessitated a translation of relevant data. In the translation process, there is a danger of meaning and context getting lost in the translation. A concept in the data can have one meaning in Norwegian, and a different meaning in the English language. In the same way, an English concept can have slightly different meaning in different English-speaking context, so it is impossible to fully avoid such misunderstandings. This necessitates a description of the research context, so that other researchers or readers of this thesis can evaluate the material based on contextual factors. I am a native Norwegian but have lived in an English-speaking country for three years, which has given me confidence in the translation process.

3.5 Research analysis

The research questions themselves are not enough to structure and limit the analysis in terms of keeping it concise, relevant and focused. It is possible to start with the research questions and write hundreds of pages about assessment, learning and challenges to these. Therefore, an analytical framework was adopted, in this thesis represented by Black and William's theoretical framework (2009) on formative assessment and William and Thompson's (2007) five strategies of formative assessment. These provides clear guidance for what LOA strategies I will look for in the research data. The analytical framework also provides a structure for how to analyse the data, into areas such as feedback or self-evaluation.

The analysis of the research data originates in the research questions of this project, which sets the agenda for what I will look for in the data. The analysis of qualitative research data is not as straightforward as the case with quantitative data. While the latter contains rigid rules and guidelines, qualitative analysis is a lot more diverse and less standardised (Bryman, 2016, p. 570). The primary goal of this thesis' analysis is to structure the research data into themes relevant to the research purpose and to filter out irrelevant data. The approach is regarded as a thematic analysis. Thematic analysis is not an identifiable approach, and it may take many

different forms from project to project (p. 585). The key feature here is to organise the data into themes related to the research questions and analyse the data thematically. Looking for repetitions in the data is according to Bryman (p. 586) one of the most common ways of organising themes. However, as Bryman notes, this is not enough. The researcher needs to argue why the themes are important and relevant to the research questions. The analysis of this research was aided by NVivo, which has helped me organise the research data into codes. These codes have allowed me to analyse the data topically and filter out less relevant data. One issue with analysis in terms of coding is that the data become fragmented into smaller concepts, which may be used out of its original context (Bryman, 2016, p. 583). Small pieces of data can be decontextualized and deviate from the original intention of the research interviewee, hence it is important to be aware of this.

3.6 Trustworthiness and authenticity

As a response to criticism to the quality concepts reliability and validity in qualitative research, Guba and Lincoln (1985, 1994, as cited in Bryman, 2016, p. 384) developed the concepts trustworthiness and authenticity as quality measures in qualitative research. While reliability and validity are associated with quantitative research where reliable and valid measurements are important for the quality of the research, this is not really applicable to qualitative research. Social contexts as in a semi-structured research interview cannot necessarily be replicated, because the social world changes all the time (Bryman, 2016, p. 383). Guba and Lincoln's concepts take roughly the same criteria roles as reliability and validity but adapted to a qualitative research context.

3.6.1 Trustworthiness

For the research project to be trustworthy, it needs a strong link between the goals of the project and the means that have been used to get to those goals. The methods must be relevant and sensible for what the project seeks to expose. This entails that all data is sampled, collected, transcribed and analysed in a similar way across all interviews, that ensures reliability between the data. It also entails that data is not changed or lost in any way throughout the research process and it needs to be clear how and why strategies have been chosen and how conclusions have been made. The research's trustworthiness consists of four related criteria, (1) credibility, (2) transferability, (3) dependability and (4) confirmability (Guba & Lincoln, 1985, 1994, as cited in Bryman, 2016).

Firstly, the research needs to be credible and it must be clear the research is conducted according to good practice (Bryman, 2016, p. 384). I have attempted to describe the research process and openly as possible, in order to ensure the credibility of the stages. Secondly, the research needs to be transferable. This criterion must be seen in the light of the critique to the replicability criterion, that research conducted on the social world cannot necessarily be replicated. Instead, it is necessary to provide ‘thick descriptions’ (Geertz, 1973, as cited in Bryman, 2016, p. 384), where the research context is described and explained, so that other researchers can fully understand the research context. To accomplish this, I have emphasised using translated quotes from the teachers, to show clearly what statements the evidence derives from. Thirdly, the research needs to be dependable. The dependability criterion stresses the importance of the research to keep an ‘audit trail’ with records of the research process, so that it is clear and obvious how the inferences and conclusions have been made. For this thesis, complete records of the sampling, interviews, transcriptions, research data, thesis and all revisions of the material have been saved, so that I can go back and doublecheck details without any information being lost in the process. Finally, confirmability stresses that the researcher must show that he stays as objective as possible. It needs to be clear that the researcher has left out personal beliefs throughout the research process (Bryman, 2016, p. 386). This may skew the whole research into making decisions throughout the research process that steers the results in a specific direction. I have tried as much as possible to explain the decisions made throughout the thesis, as well as reflecting on my own background when drawing conclusions.

3.6.2 Authenticity

When it comes to the authenticity criterion, which takes the place of ‘validity’ in quantitative research, Lincoln and Guba provide five different criteria for authenticity in the research, (1) fairness, (2) ontological authenticity, (3) educative authenticity, (4) catalytic authenticity and (5) tactical authenticity (Guba & Lincoln, 1985, 1994, as cited in Bryman, 2016, p. 386). The first criterion, fairness, relates to whether the research represents the views and opinions of the research participants in a fair way. For example, the research should not overemphasise one participant’s opinions in the research so that it skews the findings in a non-representative way. I have tried to avoid this by ensuring that I highlight differing opinions among the participants on research questions. When there have been big variances in responses to a research question, I have highlighted this in the presentation of the results.

Secondly, ontological authenticity asks whether the research improves the understanding of the social context. It needs to be clear that the research question, the research design and the methods, as well as the analysis improves our understanding of the social world. For example, if the only research question was to investigate teacher practices, would research interviews be the best approach, or would observations and an ethnographic design be better able to illuminate the research questions? Similarly, if the research's only focus was the challenges to LOA, it perhaps should have looked at document analysis or interviews with school leaders and policymakers as well. The research methods need to fit with the research questions. The third criterion, educative authenticity, asks whether I as a researcher learn about and understand better the perspectives of the research participants and that the research shows this. For example, is the researcher asking questions that he already knows the answer to? Is the interview bringing forward information that serves the research questions, or are the researcher and the interviewee discussing the subject without digging in the practices and the challenges? Prior to the data collection, I have across many weeks typed down questions I have about the topics and picked out the most relevant questions.

Catalytic authenticity, the fourth criterion, deals with the ability of the research to work as an impetus to action. Does the research provide an improved understanding that can aid us in improving the status quo? Can we say that the research is able to generate findings that can guide teachers or policymakers in a better direction? This criterion has an ambitious aspect for a master's thesis, but still, an aim for this thesis is to provide some knowledge contribution that can improve the world, or at least our understanding of it. The final criterion is tactical authenticity and asks whether the research has provided the members of the research the capacity to follow-up on the research and take action. Has the whole research process generated competence and willingness for the researcher to improve the situation of the research topic? Again, this is ambitious for a master's thesis, but nevertheless, as a researcher, I have chosen the topic strategically in the way that I want to use learning from the findings and the process to make a difference in the field of physical activity and education later in my career.

3.7 Ethical considerations

Ethics are important when conducting a research project, particularly since we document the personal opinions and voices of humans and particularly in the modern sphere of electronic devices and hacking. Bryman states that (2016, p. 123) ethical issues in research are not

completely solvable, rather one must be aware of the ethical principles and concerns. There are four ethical questions that need to be considered. These are (1) are there potential harm to participants? (2) Is there is a lack of informed consent? (3) is there an invasion in privacy? And (4) whether there is deception in the research or research process (Crandall, 1978, as cited in Bryman, 2012, p. 125).

With the first ethical consideration, harm to others, this is linked to potential loss of personal data. This project has been conducted in the standards of the Norwegian Centre for Research Data (NSD), to protect the rights and safety of both the participants in the project as well as of the research itself. The project was applied for and approved by NSD to be in line with research ethical standards. All research data was collected through a standard dictaphone, borrowed from UiO, and UiO's mobile application 'Diktafon' which is set up according to NSD standards. The application records interviews and uploads it to UiO storage service for research data (Nettskjema), and the data was never saved on the phone. The 'Diktafon' mobile application worked as a backup, to ensure that I would get away from the interviews with useful research data in case technology would crash. Recordings from the voice recorder was saved on an encrypted USB-disc until August 31, 2020. Personal data such as the recordings or files that could endanger the anonymity of the participants have never been stored on personal computers or in any way that is accessible for others than the researcher.

The second ethical consideration, lack of informed consent, deals with whether research participants are aware of what participation entails, as well as their rights as participants. Before the research process started, a consent form was made based on the standard consent form provided by NSD. This consent forms describes the research, the purpose of the research, research questions, responsible parts, reasons for why he/she is asked to participate, rights of the participant, how personal data will be handled and a signature line for the participant's consent. The consent form has ensured that the participants have been informed about their rights, and I have also taken the time to discuss the rights, the data handling and acts for anonymity with the participants. None of the research participants had issues regarding this or wished to withdraw from the research.

The third ethical consideration deals with the invasion of privacy. In this research, as participants have already consented to being interviewed, potential issues regarding privacy is linked to anonymity and confidentiality (Bryman, 2016, p. 133). We have already discussed measured to ensure this, and none of the research data are saved with information about

names, age or location with it, nor are names mentioned in the research data. The most 'private' questions in the research interviews are related to age, education background and sporting background. This may represent ethical concerns regarding recognisability. To avoid such concerns, names and names of cities have been replaced with pseudonyms, while other recognisable factors (e.g. they may have a background from a small sport) have been left out to avoid participants' anonymity being compromised.

The final consideration takes up deception as a potential ethical concern. As a researcher, I have attempted to be as clear as possible about the research's intentions and the research questions it tries to answer. It has not been in the interest of this research project not to be honest about its purpose and nature. And as Bryman (2016, p. 133) notes, more than not being nice, it would cause much harm to the recognition and professionalism of social science research if researchers attempted to deceive potential participants.

4 Presentation of Findings and Discussion

In this section, the research findings will be presented, structured according to the research questions, with subthemes for each of the research questions. The analysis will be made using an inductive approach, with a basis in the research questions and the theoretical framework on formative assessment. It is my interpretation of the interview data as a researcher that form the basis for this analysis. I have included quotations that are useful in illuminating the research questions for this thesis. I have done my best to avoid decontextualising quotes from the teachers and no sentences have been analysed without the context of their full response to the interview question.

4.1 Presentation of the teachers

Below, I present each PE teacher in this study in a table, with information on their responses to background questions about gender, age, teaching experience, sporting background and the geographical context of their school. As evident from the presentation, there are an equal number of males and females and there are teachers from a wide range of age categories. Ten teachers work in public schools and two work in private schools, one sports-oriented and one academically oriented. Most teachers have backgrounds from sports though and many of them have been competitive. All teacher names have been substituted with pseudonyms and information that may make them recognisable have been anonymised.

Table 2: showing the demographic characteristics of the participating teachers.

Teacher ID	Gender	Age	Teaching experience	Sporting background	Type of school
1 Thorleif:	Male	35-40	10+ years	Competitive in individual sports	Public, urban, medium-sized city
2 Bjørg	Female	55-60	20+ years	Competitive in individual sports	Public, urban, medium-sized city
3 Arnfinn	Male	40-45	10+ years	Competitive in team sports	Private (sports), urban, small city

4 Bjørg	Female	25-30	3 years	Competitive in team sports	Private (academic), urban, large city
5 Stine	Female	30-35	8 years	Competitive in individual sports	Public, rural, small town
6 Hallgeir	Male	35-40	6 years	Competitive in team sports	Public, urban, medium-sized city
7 Bente	Female	30-35	7 years	Non-competitive sports	Public, rural, small town (same school as Stine)
8 Tormod	Male	25-30	2 years	Competitive in team sports	Public, urban, large city
9 Marit	Female	55-60	25+ years	Competitive in individual sports	Public, rural, small town
10 Harald	Male	60-65	35+ years	Competitive in individual sports	Public, rural, small town
11 Tora	Female	35-40	8 years	Non-competitive	Public, rural, small town (same school as Harald)
12 Johann	Male	40-45	15+ years	Competitive in individual and team sports	Public, rural, small town

4.2 Research question 1: What forms of assessment practices do PE teachers use, and to what extent do they align with learning-oriented assessment practices?

In the massive amounts of research data from the interviews, I have highlighted interesting comments or responses from the teachers, that show their practices and views and that are interesting to illuminate the research question. More could be included, however, the priority has been to bring forward the diversity of the practices and to illustrate a point regarding the teachers practices in relation to the literature. Below is a table that exemplifies this.

Table 3: themes and sample extracts for research question 1.

Research Question	Theme	Sample extract
Assessment practices	Clarifying and sharing learning intentions and criteria for success	<i>These are the competence goals, now we are having ultimate frisbee, this is what we expect from you".</i> Hallgeir
	Engineering effective classroom discussions, questions, and learning tasks that elicit evidence of learning	<i>"we have four sessions with basketball, and every time they are assessed by effort and attitudes, then the last session I also look for skills. So, they had the first three sessions to practice."</i> Bente
	Providing feedback that moves learners forward	<i>"they get an assessment after each topic, but they get some sort of assessment after each session, where we try to ask what could have done different in this and this case for example".</i> Thorleif
	Activating pupils as instructional resources for one another	<i>"I want each pupil to decide a topic they want to do, they have half an hour in a session, where they can decide, or they can help in different drills..."</i> Bjørg

Activating pupils as owners of their own learning *“we’ve had a lot of self-assessment on the pupils. We try to do that after each topic we’ve had. Cause then we have made up an opinion on the assessment... Then we give the pupils the self-assessment forms, to see how it aligns with our opinion.”* **Hallgeir**

4.2.1 Clarifying and sharing learning intentions and criteria for success

One weakness to be highlighted for this section, is that not all interviews contained questions about the clarification of the learning criteria for the pupils. This is because the interview guide was made before the theoretical framework was adopted, and therefore many of the interviews lack information about this formative strategy. However, there still is evidence about it in the data.

Stine is quite clear about this and she states that she:

“present(s) the assessment criteria, repeat them every time, talk about it at the end of each session”.

Tormod on the other hand admits that the criteria perhaps are not presented clear enough to the pupils. He suggests that he talks about the criteria, but not in a formal way, whereby the criteria might not be presented clear enough. He says:

“I made some criteria for the different topics, but that I haven’t presented in a particular way to the pupils, so I think that perhaps should be presented a bit more clearly”.

Marit’s response indicate uncertainty regarding the topic. She explains that they:

“attempt to explain it to them”

Harald, on the other hand, is clearer:

“we have a talk with the pupils the first session of the semester, so they know what they’re assessed. There are never any surprises.”

Although I must admit the limitations in terms of not all interviews containing a specific question about the sharing of learning criteria, many of the interviews still have evidence about this. Some teachers, such as Harald and Hallgeir are very clear about the learning goals to the pupils and spend time explaining it to them prior to each topic or session. Bjørg and Marit also suggest they attempt to do this every session, which tells us that their approach might not be rigid. Others, such as Tormod have not shared the criteria with the pupils and recognises that this is an area of improvement, which suggests that he is not negative towards doing this more. Literature have shown that there is great variance of the degree to which criteria and learning goals are presented to pupils (Mørken, 2010, p. 47; Leirhaug & Annerstedt, 2016, p. 622) which also appear to be the case for participants in this study.

4.2.2 Engineering effective classroom discussions, questions, and learning tasks that elicit evidence of learning

This section will similarly to the literature review deal with the assessment methods teachers' methods, such as testing and authentic assessment.

4.2.2.1 Evidence of good practice

When it comes to the second formative strategy, a couple of the teachers mention specifically the construction of learning environments, that seem to be aligned with the formative strategy by William and Thompson. They do so by facilitating play, using activities and sports that few have experience with and modifying the rules to create a more even playing field.

Sonja mentions an example of an assessment situation which resembles a 'rich task' (Leirhaug & Annerstedt, 2016, p. 627; Macphail & Halbert, 2010), with a complex task given to the pupils, where the teacher assesses the process and the end result. Her example shows how PE can be a subject where complex tasks can provide basis for assessment, not only the physical performances:

"They should practice things in a way and try to reach a goal. For example, in gymnastics, I often give them a task, where they are going to put together a math task and put together (gymnastics) elements... And then you see the process they do towards the end result".

Arnfinn's case is related to performance assessment, as discussed in the literature review (Engvik, 2010, p. 209), where the pupil is not made aware of an assessment situation, but

where the teacher sets up practice sessions and make sure that he notes the ‘best’ performance of the pupil, without him or her knowing about it. Arnfinn says:

“you don’t need to signal beforehand that we are having a gymnastics test. You register what they do during the practising, right. Alright, now you can do a cartwheel, then I check that in my book... You can (also) use it as a basis for adjusting sessions, right, because it is customised education. That pupil actually needs to practice his roll more... Okay, then we facilitate that”.

Arnfinn thinks this practice saves him of time, because he is conducting the assessment during the practice of activities. Arnfinn also uses performance assessment to adapt sessions to the level of individual pupils:

Björg gives us evidence of levelling the ‘playing’ field:

“If we take ball sports first: I will always use some form of play in the game. Not just technique and game, but some sort of play, because it’s about the cooperation and get people to make an effort. Everyone can play a lot, even if they are not so skilful... If you have never played handball, it is very hard to master the activity after just three PE sessions, so facilitate for activities that not so many have done before. I think that is important in order to see everyone”.

The ‘play’ aspect of the subject is here used to create a learning environment that is less centred on performance and the sphere we recognise as PE-as-sport-techniques. Björg also mentions the example of using activities or sports that are new to the pupils, to give everyone the opportunity to experience mastery and to avoid the dominance of certain pupils:

Harald mentions an example of constructing a situation that facilitates learning for his pupils, by modifying the rules of the game, in line with Kirk’s concept TGFU. By modifying the rules of the game, Harald attempts to even the playing field and making sure that everyone can participate, and everyone have the opportunity to learn the activity:

“we have some that are very skilful individually... then you might need to say as a teacher that only girls are allowed to score, or you have to implement some rule in the game who get them to realise that they have been bad team players...”

The evidence in this section shows that many teachers are aware of problematic aspects of

assessment in the subject and attempt to solve this by engaging pupils in rich tasks (Leirhaug & Annerstedt, 2016, p. 627; Macphail & Halbert, 2010), highlighting practice of tasks and using what Engvik describes as performance assessment (2010, p. 209), where the teacher takes notes of the best the pupil can do during practice, rather than testing at a single point in time or observes the pupils randomly and potentially at a time where the pupil is not performing at his/her best. Arnfinn has much time to do this, coach and instruct pupils and catch them at their best, while teachers in regular public schools may not have this opportunity as they do not see their pupils as much. Bente handles this by using a structure that facilitates practice, where she divides the topic into four sessions. The first three sessions they practice and focus on learning the activity and the skills, before she observes them the last session. This somewhat resemble a testing environment for that final session compared with performance assessment, as described by Arnfinn, while there still is a strong emphasis on practice. Some teachers also take strategies to levelling the playing field, in order to avoid the subject favouring pupils that do sport and that are competitive (Andrews & Johansen, 2005; Wiken, 2011; Seelen, 2012, p. 200; Säfvenbom et al, 2014; Leirhaug & MacPhail, 2015, p. 635; Kolbeinsen, 2016). This occurs through using play, rather than traditional sporting competitions, and modifying the rules of the game in line with TGFU (Kirk, 2010) which reduces exclusive behaviour by those that have better technical abilities in for example ball sports. Although the evidence come from only a few of the teachers, we can clearly see that many teachers are aware of assessment-related challenges and take steps towards more learning-oriented practices, in terms of the second formative strategy by William and Thompson (2007). However, the lack of such evidence from other teachers suggest that this is an area which requires improvement across Norwegian PE classes.

4.2.2.2 Testing

In the literature, testing comes up as a form of assessment that does not facilitate much for learning, but rather as a technical tool to rank pupils, which is incapable of evaluating broader learning goals (Kirk, 2010). In this study, the findings indicate that this is a practice most teachers attempt to steer away from. However, some testing is apparent.

Thorleif talks about cooper tests, but notes explicitly that they are not using testing for grading purposes, except in swimming, which they outsource to a swimming coach:

“she has quite clear boundaries for the different grades. For example, to be able to get the grade 6... you should be under 2 minutes on a 100m. I have asked whether it is

okay to do it that way, but she says it's a good indicator on your swimming ability".

Sonja notes that she takes the time during running sessions but that it is not something she uses for assessment. She states that she does not like it and she say it is not how they are supposed to use the test, but that she feels she has no choice in appeal cases:

"I must admit I use it in grade appeals... Here you can see what the average time for this course is".

Bjørg is not at all fond of testing in PE and is neither using tests for summative or formative purposes:

"I don't do tests. At all. We can go for a run, but then it's everyone at their own pace, wherever they want, so it's not that visible who comes back first".

Stine describes the use of summative tests for formative purposes:

"we don't have any concrete tests, except for a local round... We run it twice a year. When I went here, we got grades based upon our time. That's not the case anymore... At the same time, it's hardly doable to get a high degree of goal attainment if you clock in at above 30 minutes".

Hallgeir claims that many teachers he knows of uses testing, but he is avoiding it:

"very little testing. We don't have much basis for doing it either".

Harald, who as over three decades of experience as a teacher suggests there has been a shift in recent years when it comes to testing. He explains that testing now has a much more formative role:

"we were much more concerned about performance then as opposed to what is now more and more relevant, and we steer more and more away from testing... now tests are used completely different in that it is some sort of progress assessment".

Johann feels that the grade in skill tracks mostly reflects their ability level in the game. He also feels that it is very hard to assess everyone during a game play, hence the skill track gives him the opportunity to assess each of them individually. Johann is conducting testing in the traditional way but recognises that he is not supposed to assess them based upon their

performance in such tests. He finds it challenging to know how to assess such activities:

“when I’m going to give a grade, I set up a course and film one by one as they go through the techniques they are supposed to... the easiest to assess is athletics. There the criteria are already set (in the local learning plan), so I can just take the time... The hard thing is that you are not allowed to assess them based upon the time they perform”.

While much of the literature suggest testing culture dominates the field of PE (Lopez-Pastor et al, 2012; Græsholt, 2011, p. 104), the evidence from the interviews largely suggests that teachers are very concerned about testing and has moved away from it, which also experienced teachers, such as Harald, highlight. Younger teachers in the study are very concerned about testing in PE and suggest that this still is the basis for assessment in many classes. Hallgeir and Tormod emphasises that older teachers in their schools are still using physical testing as a basis for assessment, which may suggest that testing is still prevalent and that teachers that use such practices may have avoided participation in this project. Although most teachers in this thesis refrain from grading tests, many use tests for formative purposes (Leirhaug & Macphail, 2015, p. 630), where running tests are conducted twice per year to assess their progress. Johann’s use of skill tracks represents one type of testing that fit with the behaviourist paradigm of PE, that Kirk (2010) suggests is kept within PE due to its ease of measuring ability. The evidence suggests that testing still has a place in PE practice, but that the grading of test results seems to be on a downward trend. There are however limitations to this study, in that many teachers that may still do this might refrain from participating in the study.

4.2.2.3 Authentic assessment

Observation as an authentic assessment tool

In the literature review, authentic assessment was mentioned as one of the main assessment strategies in PE. In this section we see a few examples that highlight observation as a main tool, where pupils perform an activity (authentic assessment) whereby the teacher observe and assess. The examples below show evidence of this assessment strategy.

Arnfinn uses observation as the primary tool of assessment:

“if you take football, the topic for one session can be passing and receiving the ball,

then I assess that, through observation primarily... of course, you can bring a sheet which you take notes on... but that's more rigid in a sense. Much of the running assessment happen through observation”.

Tormod’s statements show that he uses his assessment sheet as a basis for his assessment, where observation represent the main tool:

“most of it is an observation of the pupils, so I have different sessions, perhaps double sessions, or three weeks with the same activity... Then some of the observation is that I have a sheet for every pupil, about effort, fair play, skills/knowledge, where I set pluses and minuses to create some sort of grade...”.

Marit confirms what the others say:

“yes, observation is the main assessment form really”.

The findings are in line with Jonskås’s (2009) and Hansen’s (2017) studies which found observation as the main assessment tool used by PE teachers. Observation and authentic assessment represent something positive in terms of LOA in terms of letting pupils conduct real contextual tasks, rather than decontextualized testing (Engvik, 2010, p. 109), however, there are many pitfalls to this, such as whether the pupils are experiencing the sessions as a learning environment or a competition environment. Another issue, which we will explore more in the section on challenges, is the background position the teacher may take during observation, in order to document and evaluate pupils’ performances, rather than take active part in the session and coach and involve pupils.

Taking notes

Notetaking as a tool for documentation about pupils’ learning and performance has come up as a key tool in the assessment process in this study.

Thorleif is not fond of documenting much during sessions, but feels he have to, and he notes some troubling aspects of notetaking:

“some finds it uncomfortable that we take notes, but we have many pupils, so if we are going to have a chance to remember, so sometimes need to use the notebook and write some pluses and minuses... at the same time, I don’t like noting down too much,

because it's nice to be present in the class, to have contact with them and not just stand with the notebook... the optimal thing would perhaps be if I could remember it all without taking any notes during classes."

Stine takes notes of each pupil each session, but does it after the session:

"the way I do it is that I always bring a sheet in the sports hall... then I write down shortly after the session what impression I got of the pupil, say in basketball, whether it is low, medium or high degree of goal attainment."

Harald says it is challenging to get an impression of everyone, hence notes are important. He also mentions that notetaking is dependent on their time availability:

"it's rare you manage to get an impression of everyone, but as much as possible, we try to take notes after the class, on whether one or the other has remarked himself or herself in something particular... we have some PE sessions there and swimming sessions there, you have enough of time pressure changing clothes and getting here and there, so it's not always we get to make so many notes, so then it's a matter of remembering."

It is apparent that notetaking is an important strategy for the teachers in terms of gathering enough evidence for the grading process, which is linked to summative rather than formative purposes, which represent a potential threat to the LOA in sessions. The negativity associated with testing and the need for documentation due to grade appeals and regulations may have led to documentation of authentic situations taking over for testing as an important assessment method. The documentation of pupils with regards to the competence goals is also in line with literature about MBO and NPM (Møller & Skedsmo, 2013). Thorleif notes, the optimal situation would be if you could be fully present in the activity. In addition to this, he suggests that some pupils find it uncomfortable, which may lead to avoidance strategies (Andrews & Johansen, 2005). Stine's approach is more compatible with LOA, as she does the notetaking after sessions, however, as Harald notes, time after the class is rather restricted and it can be hard to remember pupil performances if it takes too long before teachers can make notes, which may lead to some teachers taking notes during sessions and eating of the time they have to facilitate learning. Literature has showed us that teachers find it hard to gather evidence about pupils' abilities and with the decreasing popularity of testing, we may see teachers taking a background position in order to takes notes (Arnesen et al, 2013). This

was also evident in the interviews, where Bente regularly stands on the sideline, taking notes to document pupils' performances.

The use of video

The use of video in PE assessment is briefly mentioned in the literature and also comes forward from interviews as a method used by many.

Thorleif feels that filming the pupils give him more freedom to instruct:

“now they (the pupils) have an Ipad and for some topics, such as dance or gymnastics they have group tasks, which they film and deliver. Then we can put away the notebook, walk around and help them and stuff, then we can sit down and assess (the video)”.

Arnfinn brings up both that video as a tool that allows him to be present and as a tool to coach pupils technically:

“it's all about being present and observe, and you can use video. Experience from that is also useful for the pupils. Because if I say no we're doing a cartwheel... And I say, your legs are bent, (and they say) no. No? Come and watch yourself”.

Stine feels that video is a necessary assessment tool for the grading process:

“the only time I use video is when I assess dance...Because then they are in groups of maybe six... And it's completely impossible for me to set a grade on all six when watching them dance, so then I film them.”

Johann experiences issues with assessing many participants at the same time in authentic assessment situations such as game play, which he solves through using skill tracks and film:

“you don't remember if you have (just) observed them, in a way... In practice, I'm not able to assess game play, if I don't set up a video camera which films the whole game... I set up a course containing the intended goals, then I film as they go through it, then I sit down and assess the film afterwards”.

As Thorleif and Arnfinn suggest, video may be a method that solves some of the potentially detrimental effects of notetaking, as the teachers can put away the notebook, be present and

coach pupils more than they can if they take notes, because they can watch the video later. The question is whether the video may be another stressor for the pupils, as indicated by Cassady and colleagues (2014), however, the teachers have not mentioned this as an issue. While video may facilitate the use of coaching, most of the examples from the interviews comes from dance, gymnastics and as a method to assess skill tracks in ball sports. This suggests that video is mostly a tool for technical skills in the subject. It does not seem to have a large role in assessing non-technical competence goals linked to cooperation and fair play.

4.2.3 Providing feedback that moves learners forward

Feedback constitutes the third formative strategy by William and Thompson (2007) and there is much statements in the data regarding this.

Arnfinn states that they have a coaching perspective on feedback in the private sports school:

“we don’t have a rigid structure on how assessment is conducted. Most of the assessment with us is running. It’s coaching in activity.”

Hallgeir explains he gives feedback in between the formal assessment after each semester:

“you give continuous feedback throughout both the topic and the school year. And you try to get a more serious talk to each pupil for half-yearly and final assessment.”

Bente might have the most rigorous approach when it comes to feedback in this study:

“I’m happy with myself that I bother writing written feedback to each pupil after each session. And I’m annoyed with myself sometimes for doing it, because it takes so much time.”

Tormod talks about a potential issue where eager pupils will get more feedback:

“it happens every session, some pupils are much more ‘on’ and wants to know. Then they’ll get more feedback (than others)”.

Johann realises he has a lot of potential when it comes to feedback:

“it’s clear, I have a lot to work on as a PE teacher, take them aside and maybe do the skill track twice, or do the test twice... it just takes time... That’s why you don’t do it (as much). Now we’re done with football, now we’re gonna start something else”.

The data reveals that teachers focus on giving feedback more regularly than the half-yearly assessment, however, there is great variance in how they go about it. Arnfinn from the private sports school has described they have so much time for the subject that he can focus on continuous coaching of the pupils. A few teachers, including Johann, do not feel they are able to give enough feedback, while Bente writes a written feedback for each pupil per session, which she describes takes a lot of time. This aligns with Mørken's findings (2010, p. 49), where many teachers do not feel they have time to give feedback. While Bente's approach represent perhaps the most continuous and formalised structure for feedback in this thesis, Arnfinn's case from the sports school shows close connection between the teacher and the pupils in a learning process. He has so much more time for physical activity than the other schools in the study, which explain the relaxedness regarding grade-setting, while many teachers from the public schools on the other hand feel pressured by time and are less able to spend time on feedback. Research about feedback in Norwegian PE suggests there is great variance to the degree it is used by teachers (Mørken, 2010; Græsholt, 2011; Hansen, 2017; Thorsnæs, 2019) and the degree to which pupils recognise the feedback (Leirhaug & Annerstedt, 2016). While a whole master's thesis should be done on feedback itself in PE, the evidence from the interviews confirms that there is great variance to the degree teachers are able to give feedback.

4.2.4 Activating pupils as instructional resources for one another

There is less evidence in the interviews about the activation of pupils as instructional resources for one another, such as through peer assessment.

Marit uses pupils as instructional resources:

“in ball games and gymnastics, I often use the pupils, because usually there are some that are active gymnasts... And then I use the pupils, then they get the mission to design a scheme we can use for the session.”

Of all the research data, only three teachers have examples linked up to the formative strategy. While Bjørg, as shown in the previous table, tries to involve pupils in the session, Marit activates pupils as instructors for topic they know well. The key finding is obviously that peer assessment seems to be less evident among teachers' assessment practices, which is in line with Leirhaug and Macphail's paper (2015) where only a single teacher out of 23 used peer assessment. Pupils in Thorsnæs' study (2019) seemed to have more experience with it,

than what the evidence from this thesis suggests, however, the prevalence can only be explored in a quantitative study.

4.2.5 Activating pupils as the owners of their own learning

There is much more evidence when it comes to self-assessment. However, the degree to which it is being used varies a lot.

Sonja describes an occasional practice and a ‘grade-setting’ self-assessment:

“sometimes they have self-assessment, when they start a new topic, they check out in a form what level they feel they’re at, bad, medium, or high level on a scale and at the end of the topic they do a final assessment (of themselves)”

Bjørg tells me that she has not used self-assessment and she is a bit sceptical:

“I haven’t practiced that yet. I find it very challenging, because I notice that many pupils are not really aware of what PE is. They believe PE is about playing football...”

Tormod describes a double-edged sword when using self-assessment:

“yes, and there is varying success with that... they get very disappointed that writing is implemented in PE, so some get very short and concise, and they try to respond in single words, or short sentences, which makes the self-assessment lose some value... Some are better written (than orally), or find it uncomfortable to have a face-to-face conversation, so the self-assessment has given me something”.

Tora talks about the lack of making pupils owners of their own learning:

“assessment for learning? I think... That should have been... I think that the forms we provide the pupils with, have been good, to make pupils engaged, but perhaps... Maybe they should assess themselves more throughout the year...”

Many teachers in this study have some experience with conducting self-assessment among pupils, while some have no experience with it. Johann, Stine and Tora feel they use it too little, while Bente has been hindered by the change of online learning platform. These results align with what pupils responded to UDIR’s questionnaire in 2017, where only one fifth of

them felt they took part in their assessment (Utdanningsdirektoratet, 2018). For the teachers that uses self-assessment as a method, such as Bente and Hallgeir, it appears from their responses that they are preoccupied with how it aligns with their own evaluation. This aligns with literature in that it indicates self-assessment being used more as a tool for ‘reality-orientation’ (Græsholt, 2011, p. 29), rather than a tool to engage pupils in deeper learning, although some teachers may emphasise this as well. In such instances where self-assessment primarily consists of writing down what grade you feel you deserve in the subject, I question whether it pushes the pupil to become owners of their own learning (Black & William 2009, p.7), as pupils merely seem to guess what grade they will get (Leirhaug & MacPhail, 2015, p. 636). Tormod’s example show that the latter might be the case for many, where pupils provide short responses in self-assessment and are unmotivated for written self-evaluations. Although it might be normal for pupils to be unmotivated to written work in a physical subject, it hardly suggests that the pupils have taken the roles as owners of their own learning. Together with those that seem to use self-assessment as a ‘reality-check’ tool and those that do not use self-assessment at all, evidence suggests that the practices are far from the definition of the fifth formative strategy.

4.3 Research question 2: are there barriers to teachers using learning-oriented assessment practices in their PE practice, and what kinds of support or advice do physical education teachers need to develop educationally sound, successful and sustainable forms of assessment?

Much of the literature discusses issues that have implications for assessment and LOA. The results below show evidence of many of these issues and also highlights some others. The categorises below have been formed by common challenges identified from the research interviews. The quotes are selected in an attempt to show the breadth of the challenges.

Table 4: themes and sample extracts for research question 2.

Research Question	Theme	Sample extract
	Teachers’ view on learning in PE	<i>“I find it hard, because... We don’t have that many PE sessions, so that... We don’t do the same activity for so long, and then it can be hard to see the learning curve.” Bjørg</i>

Challenges to LOA	Student and parent expectations	<p><i>“there is no subject where you have so many to assess for you. Who suggest they can assess for you, everyone sort of knows what grade their kids should get in PE, because they are so good football players or they are fast skiers...”</i> Sonja</p>
	Documentation needs	<p><i>“I haven’t had much problems (with appeals)... But I was very busy with it in my previous job (in public school) ... Then it was very much focus on documentation in relation to appeals, because there are two subjects that get most of the appeals; the aesthetic subjects.”</i> Arnfinn</p>
	Motivation and stress among the pupils	<p><i>“(Grades) drives them. Because it’s not always they are motivated. Grades is a motivator for many to keep going... If you have some strength training and you turn your back on them, they stop doing push ups to put it that way...”</i> Thorleif</p>
	Pupils’ (lack of) participation	<p><i>“that’s been a big part of my job this fall, trying to get them motivated, and get them attending at all... Those pupils that don’t want to attend. That’s a big part of my job actually... Those that struggle with overweight for example... They rarely like the subject.”</i> Marit</p>
	Competence goals and local learning plans	<p><i>“with the sports I know personally, those I think are quite easy to facilitate... But those I haven’t personal experience with myself, those I get quite uncertain about. The competence goals in swimming... What’s good crawl technique? How do I judge breathing technique?”</i> Johann</p>
	Teacher education	<p><i>“uhm we went through some assessment methods, but that’s an area I’ve felt uncertainty about and still feel uncertain about. I try to look through various online resources to try to get to the bottom of it.”</i> Tormod</p>

Resources for teachers	<i>“most of what (online resources) I find, is stuff that teachers have made and posted. I don’t feel that... I don’t think the stuff online from UDIR is good enough for example...” Bjørg</i>
Class size and time to assess	<i>“we have big classes now, 29-30 pupils. You can’t assess every pupil each session. We’re not able to do that. Then they won’t learn anything. That’s the issue...” Thorleif</i>
Facilities and resources	<i>“the space is... very little. One of the classes always have just half a handball pitch, and that’s in a way okay, but I have PE tomorrow and then we’re three eight grade classes there at the same time, and it’s obvious that there is not much space, so you have to adjust to it...” Bente</i>
Private vs public schools	<i>“we have quite high competence on our teachers and coaches here, so our demands are high, but we have eager pupils you know, they have chosen to go here, paid to go here, so it’s a completely different setting than a public school, where half of the pupils cannot be bothered.” Arnfinn</i>

4.3.1 Teachers’ view on learning in PE

In order to map out the terrain in terms of LOA, the interviews took up questions about learning in PE and its relation to assessment. Teachers were asked about whether they thought learning occurred in PE and whether assessment and grades reflect learning, or whether grades merely reflect the pupils’ status quo levels in activities.

Thorleif expresses his concern with assessment in relation to lifelong learning:

“we have discussed many times that the most important is that they are active later in life, that’s the most important thing, but I don’t know how good it is to assess them all the time in relation to that...”

Arnfinn highlights their focus on learning and the limited role of grades:

“we have the main focus that the education is a learning course, meaning that we

practice... It's first when you get halfway through tenth grade that you are actually using the grades for something...

Stine do not think there is enough time for learning:

“not enough (learning). Not like we were taught up in my studies. Because... Firstly, PE has way too little time in relation to everything they are supposed to learn. So there should be many more hours with PE...”

Hallgeir suggests assessment eats from the time they have to practice:

“I think that if you removed the assessment, one could have gone deeper into the pupils' learning on this. Not assessing them, but giving them the opportunity to master different things... you get more time to practice and get better...”

Tormod brings up differences in abilities as a threat to learning:

“I think it's hard to know how much you can expect them to develop certain skills when you have so few opportunities to practice? The gap between the pupils is so big, and it's hard on a small space to sort of, display your skills in a good way...”

Johann spends much time testing the pupils and feel there is not enough time for learning:

“a lot of status quo... (the assessment) captures their status quo, and... I use a lot of time on the assessment, so say, we do two sessions of practicing basketball, then you have to use one session to conduct a test, let them do it one by one, then a third of the time is used on the test, which they don't learn anything from, you're just checking where they're at...”

The evidence shows quite gloomy views on learning in PE. Some teachers, such as Thorleif highlight that they see assessment as conflicting with what they are trying to achieve in the subject, while Bjørg, Stine and Marit do not feel they have enough time to see a learning curve. The literature explored goes into challenges of time in relation to providing assessment (Mørken, 2010, p. 49; MacPhail & Halbert, 2010, p. 36; Hansen, 2017, p. 48; Bezeau et al, 2020, p. 389) and the teachers' responses also add to the perceived time they have to facilitate learning. This suggests that many teachers find it hard to see how they can create learning opportunities, where they rather feel they are assessing pupils' status quo in the subject. It is

suggested that the assessment takes away time they have for instruction and that level differences inhibits the learning potential. Only one teacher responds solely positively about learning in the subject, which indicates that there are some fundamental structural issues regarding assessment and learning in PE. The way teachers see assessment as a barrier to learning in the subject suggests that assessment of learning rules over assessment for learning in the subject, and when there is not enough time to see learning progress, assessment obviously represents a big challenge in PE.

4.3.2 Student and parent expectations

The expectations and pressure teachers feel from pupils and their parents has been identified as a potential challenge to LOA in the research data. While it is obvious that this possess a challenge for assessment in general, the link to LOA is less obvious. However, one teacher talks about the pressure from competitive boys in the subject, which may create a challenge for teachers in the implementation of LOA practices. All resistance that PE teachers experience in relation to assessment may also influence their perseverance or ability to focus on a learning-oriented environment rather than an environment that focuses on grades.

Bjørg highlights pupils' lack of respect for the teacher's decision:

“that I find interesting with PE, there is no other subject where the pupil feel he has so much to say. In the Spanish lesson, the pupils come there, there is a schedule, that's all good... I don't know what they think, they have a lot of expectations that PE is supposed to be how they want it to be.”

Stine talks about the differences between younger and older pupils:

“it's big difference on whether the pupils are eight, ninth or tenth graders. The eight graders are a bit like, come from primary school, they have played a lot... So they don't miss feedback so much either, cause they have PE to have fun...”

Tormod goes into the group that gives him the most to deal with:

“boys with a high competition instinct... They struggle with respecting the way the assessment is conducted. Some pupils think that they I've just made up some assessment formula.”

Marit find that compare their assessments and complain to her about their grades:

“all pupils are not able to see themselves, right, they compare themselves a lot to friends. They are very...They can for example say that, well, he got a 5 and I got a 4 and why is that, and so on... It can be quite challenging, I think.”

Teachers talk about both parents and pupils feeling they have something to say about the how assessment is conducted, and how grades are given in the subject. Throughout the interviews it is apparent that many teachers face pupils with clear expectations that PE is an activity subject and not a learning subject, in line with what literature describes (Seelen, 2012, p. 178; Wessel, 2014, p. 82; Kolbeinsen, 2016, p. 45; Hansen, 2017, p. 36). Hallgeir and Tormod mention the lack of respect from pupils in the interview, which might be linked to the finding that pupils see PE as less important than other subjects and struggle to see the learning aspect of the subject (Wiken, 2011, p. 70). While pupil involvement is something that should be implemented more in the subject, this says something about the everyday life of the PE teacher, where teachers must deal with tiring pressure from multiple actors. This pressure, as we later will explore in the next section, has contributed to a documentation pressure due to grade appeals that challenges the potential for LOA (Jonskås, 2009, p. 81; Arnesen et al, 2013, p. 24). Hence, the pressure from pupils and parents represents an underlying issue that teachers must deal: the need to document pupils’ performances.

4.3.3 Documentation needs

The research interviews early exposed the need to talk about the need teachers experience to document pupils’ performance and abilities, in case of a grade appeals. The teachers have described that without documenting enough, pupils will likely win in appeal cases. The teachers are frustrated that the time it requires goes at the expense of instruction and learning.

Sonja suggests the session get very grade-driven:

“you get a little bit controlled in tenth grade based upon the grade appeal thing, because everything needs to be documented sincerely and that destroys quite a lot. Pupils complain on their grade, then you need to have set goals for all periods and for all topics in the learning plan, then we need to document that the pupil unfortunately does not have the skills... it all gets very grade-driven.”

Stine feels that the grade-focus steals from the time she has for teaching:

“... if it was a subject without grades, we could have focused on promoting enjoyment

with activity, but now the goal constantly is to (focus on techniques)... because the pupils have these terrible requirements hanging over them and we have to document everything all the time... So they're used to getting written feedback, and they want oral feedback and it takes up so much time that you could have spent teaching, and that's what I find very frustrating..."

Hallgeir has a lot of concern about the documentation need:

"... I sit down and assess them after each session I've had. And try to write notes on each pupil in relation to low, medium or high degree of goal attainment of the local learning criteria we have given them. Enormous amount of work... The documentation need is frightening... because there are so many grade appeals in today's school... It's been a sad development in my eyes, and it coincides with lower and lower status of the teachers, and less and less respect from the pupils and both parents and pupils... demanding more than what they have the right to and more than what they deserve. Which results in many teachers giving out better grades than what they actually can vouch for."

Bente describes the strategy she uses to facilitate documentation, where she first uses practice sessions, before she:

"the last time I just sit on the tribune, then I give them many different exercises that they have done previously, so they knew what to do, and then I take notes, for example in basketball... Then I assessed only skills, so I just sit there and watch. I do it this way for all activities."

Johann talks about the documentation need leading him to spend much time on tests:

"yes we document ourselves to death... in relation to a appeals in tenth grade... Eight and ninth grade they don't care so much, then I don't need to worry about appeals, then I discuss with the pupils, because then we have enough time. But in tenth grade we are done, then your grade is counting. And then, you need to have your documentation ready, if that appeal comes in... So I have to document it. And then you get caught up in those tests... Use a lot of time on that rather than developing skills... I feel that I... grade them based on status quo rather than... progress... we're basically documenting based on a grade we already know."

The documentation needs highlighted by the teachers suggests that much of the time, that potentially could be used for more LOA practices, goes into documenting pupils' performance to justify the grades of the pupils. This seem to cause a lot of frustration, as they need to justify "what we already know", as Johann says. The need to grade every pupil in each competence goal and in each topic seems to create a competitive environment in line with what we find in the literature (Säfvenbom et al, 2014; Leirhaug & MacPhail, 2015, p. 635). The documentation need seems to cause teachers to facilitate activities that allows them to document skill levels and performances, as Johann suggests, rather than spending the same time coaching them and facilitating learning. Both Sonja, Stine and Hallgeir specifically highlights this. While the competitive nature is PE is well covered in the literature, less literature goes into the need to document performances. Arnesen and colleagues (2013, p. 24) described how the documentation need led teachers to take a background position, where they could take notes and assess, as opposed to instructing and facilitating further learning. As shown above, we see that Bente uses this strategy, where she has practice sessions first, then taking a background position to assess. Johann solves the documentation need by using tests such as skill tracks. Arnfinn seems to be the teacher that feels pressure about this, because they have so much PE and he enjoys great trust in the sports school from both pupils and parents that are less likely to complain about grades, because they perform at high levels. The evidence from the interviews clearly indicates that the need for documentation requires a lot of time that goes at the expense of the time they have to instruct pupils and therefore LOA.

4.3.4 Motivation and stress among the pupils

Pupils' motivation represents another potential challenge for LOA. As we will see in a later section, Arnfinn, who works in the private sports school, talks about the privilege of having pupils that love sports and that want to improve, which reduces the time the teacher needs to spend trying to get pupils to attend.

Arnfinn describes pupils as preoccupied with their grades:

"the pupils stress the assessment thing all the time, they are so preoccupied with that number. But we try to focus on the process... we encourage the pupils to remained focused on the feedback, more than the number. And that's a question, whether there should be a grade or not..."

Stine describes an assessment situation where pupils display discomfort with the assessment:

“you constantly have to try and get an impression of the development... But at the same time, it leads to the pupils knowing they are constantly observed... Observed in a way where they are being assessed. And many pupils find that tiring. And it’s also a stressor for many pupils, and I think it takes away a lot of the joy I wish to give.”

Tormod attempts to avoid graded assessment up until the half-yearly assessment when the pupils are entitled to a grade:

“always at the end of the first semester, they are entitled a numbered grade. We have consciously tried to avoid numbered grades until then, so they get only written feedback, to reduce the grade pressure, but I notice both in eight and ninth grade that they are very preoccupied with... they might look a little bit at the feedback, but it’s the grade they’re after.”

The teachers above describe the graded assessment as an extrinsic motivation that according to Ryan and Deci (2020) will overwrite intrinsic motivation. Getting good grades becomes more important to pupils than learning and motivations of the pupils contribute to the creation of learning- or competition-oriented environments in PE sessions (Ingebrigtsen and Mehus, 2006). Thorleif describes a situation where many pupils will avoid putting in an effort if he does not watch them. Arnfinn and Tormod show evidence of pupils that concern about their grade. Stine describes an assessment practice at collision with the lifelong activity enjoyment objective of the subjective (Utdanningsdirektoratet, 2015), where she wants to create a joyful environment for the pupils, but where pupils feel the stress of the graded assessment.

Thorsnæs (2019) highlighted from interviews with pupils that that grades represent a pressure for pupils in PE also seem to conflict with the learning objective. The extrinsic motivation assessment seems to create is not in favour of LOA, as intrinsic motivation more likely leads to learning than extrinsic motivation (Ryan & Deci, 2020, p. 2). Hence, teachers potentially face an obstacle in LOA, as pupils are more concerned with the extrinsic motivation of grades than with the intrinsic motivation of learning

4.3.5 Pupils’ (lack of) participation

From the interviews, pupils’ participation arises as a challenge that many teachers need to devote time to. Many pupils avoid participation in the subject for various reasons, which creates a difficult environment for the teachers to facilitate learning.

Thorleif talk about pupils that are not attending:

“firstly, an optimal assessment situation is where all pupils are attending. That you don’t have to use a lot of time on pupils that find it hard to shower, hard to attend, that are hurting here and there. You use a lot of time and lose a lot of the overview because of those things.”

Stine talks about certain groups that struggle with participation:

“it’s very hard to get muslim girls to attend swimming class... And we have those pupils that are generally demotivated for school, they are not always motivated for PE either... As with the wardrobe situation they feel it’s best to just avoid the whole thing, maybe they feel a bit overweight, and instead of risking getting bullied, they just avoid the whole situation.”

Johann is frustrated that he is not able to involve those that do not want to attend:

“then you have those that ‘forget’ PE clothing... Right, they are the hardest ones to deal with... I feel guilty for them, because they are the ones that need PE the most. It’s not hard to understand them, they hate PE... And to turn around that, the psychological, because... We don’t know what the causes for that are. What is it that keep you away from being active? Because that’s not normal. The normal thing for kids is to be out and play.”

Although it must be said that many teachers, such as Tora, says that most pupils attend, there are still some pupils that do not want to attend. This clearly represent a challenge for LOA for those pupils that do not attend, however, it also eats of the time to the teachers that have to spend time trying to motivate pupils rather than instructing. Literature confirms the issue that many pupils do not want to attend (Andrews & Johansen, 2005; Ommundsen & Kval, 2007) while little literature looks at the effects this have for teachers. What Thorleif describes emphasises the varied issues that PE teachers must deal with during PE sessions, which may inhibit their opportunity to develop their assessment practices. While many may aspire of providing pupils with meaningful feedback and engineering good learning environments, there seems to be a constant ‘noise’ in the subject of the various issues Thorleif describes. The struggle that teachers seem to have getting everyone attending may not only a barrier to designing educationally sound assessment in the subject, but also a symptom of the lack of

such an assessment system. As one of the teachers, Johann, describes: “the normal is that the kids should be out and play”. If LOA was already implemented and activities were inclusive rather than exclusive, the lack of attendance may be less of an issue. However, in defence of the teachers in this study, they say that most pupils do attend, while those that do not attend causes frustration with the teachers and requires much time.

4.3.6 Competence goals and local learning plans

In the interviews, there is a lot of talk about competence goals and the difficulty many teachers experience in interpreting them and using in assessment. Much of this related to the reliability of grades in PE and is beyond the scope of this thesis. However, some evidence is linked to challenges teachers face in relation to LOA.

Sonja feel that the local learning plan highlights technical aspects of the subject:

“we have a local learning plan in PE that we use in our school and period plans and year plans and within each period plan we have goal criteria which steers us... And... I don't like that very much, to sort of, assess whether you are able to do a headstand, if you can walk on your hands, do you have bent legs, right... That's not what I feel should form the subject.”

Bente finds the goals and guidelines very flexible which she enjoys:

“uhm, they are very open and... They are so open that you can have many types of activities within the same goal... So I think it works good the way it is now.”

Marit thinks there are too many competence goals:

“there are so many competence goals. I find it hard to get everything in... It's something about these learning plan that are... Well, there is so much in it, so many competence goals in things, so it's hard to be able to see everything, you know.”

Sonja highlights a central barrier to LOA, where the local learning plan, which is based on the competence goals, facilitates a technical view on ability in PE. This has implications for the way Sonja teaches and assesses, in a way that strengthens the need for documentation of specific skills, rather than spending more time focusing on learning activities. While Bente enjoys the flexibility the goals provide, Hallgeir agrees with some of the literature in that he found the goals too vague (Jonkås, 2009, p. 71; Mørken, 2010, p. 52). He and Tormod desire

more clear guidelines which is supported by the literature (Annerstedt & Larsson, 2010, p. 107; Hansen, 2017, p. 46). Marit thinks there are too many competence goals to assess. If competence goals are too many for teachers to be able to assess, this may have implications for their need to document during sessions. While many goals and many aspects of the subject may create opportunities for pupils to try out different things, it also creates what Kirk (2010) describes as a molecularised subject with many learning goals to go through, with limited time for deeper learning. In this sense, what Marit describes together with evidence about PE-as-sport-techniques may suggest a subject that facilitates shallow learning. Johann feels he has no issues in teaching or assessing activities that he knows well himself, however, he struggles with instruction and assessment in activities where he lacks knowledge. Literature suggests that PE teacher often emphasise their own activities in the subject (Kirk, 2010, p. 63; Moen, 2011, p. 234) and the lack of teacher competence in relation to competence goals represent a barrier to LOA.

4.3.7 Teacher Education

With regards to education about assessment in their teacher education, the participants were asked about what they learnt during their education and how it has helped them as teachers.

Thorleif were surprised of how little they talked about assessment in his education:

“we have talked a bit about it, but it was surprising how little talk there was about assessment, about what should be weighted and not weighted, but I’m going to a seminar about the new curriculum today, so hopefully there’ll be something about assessment... But when I took the PPU (practical pedagogical education) there was little about assessment. It was very superficial.”

Stine had a lot about assessment, but feels she has less time in the real world:

“there was a lot of focus on it, particularly in my bachelor... We discussed it definitely at master’s level as well... But you don’t have as much time in real life as you have when you study... So that’s a big difference. Because if you had all the time you need, you could have taken the time to give good and constructive feedback continuously...”

Bente missed more practically oriented assessment during her education:

“yes, I think there was, but... not that you can use in school, there should have been a bit more practical assessment perhaps. We had practice, right, but the way we were taught it there... I think it’s a bit different in practice here than what you learn during PETE. You learn about a lot of things, but practical stuff, such as a normal day for a teacher, you don’t learn about that.”

Johann describes it as a shock when he felt how assessment was out ‘in the field’:

“uhm... Way too little. Too little practical assessment. I’ve talked with those I went to teacher education with, and... it was a shock for us all when we were thrown out in the field and started assessing... We weren’t trained properly in it...”

There is close to unison among the teachers about the limitations of what they learnt during teacher education. Based on the teachers’ comments, there appears to be a discrepancy between an ‘ideal’ assessment scenario during education and the reality that they face every day in schools. Both Kirk (2010, p. 47), Mørken (2010, p. 49) and Engvik (2010) highlights formative strategies such as feedback puts pressure on their professional competence, and in the light of evidence about documentation needs, teachers seem to face an overwhelming reality in terms of teaching and assessment in the subject following their education. The question regarding LOA, is to what degree this is taught during teacher education after the implementation of AFL in the education system (Utdanningsdirektoratet, 2018) and whether teachers are prepared to implement these in real life. This could be a master thesis itself.

4.3.8 Resources for teachers

When it comes to teacher resources, such as seminars, online material and resources, some teachers suggest that they miss this in PE.

Bente talks about the lack of opportunity to attend seminars because of time barriers:

“I know that upper secondary school have these seminars... where they have all the teachers working with sports and PE, perhaps in the whole region, gathering once a year and get new input... But not all of us PE teachers can go, because someone needs to keep the wheels going here. There are many PE teachers that have other places to be right, so even if you want to learn more about PE, you know that if you’re going to spend one day away from your class, you need to get in a substitute and do all those arrangements, so that’s a lot of work.”

Tormod feels he misses a textbook in PE and struggles to find online resources, which he feels he need as a newly educated teacher:

“Why not create a common idea bank over different activities, here you have a range of different sports and activities, that you can use... So I’m trying to find online resources, and I think some competence goals could have been more specific, or that you had more resources to creating good sessions for the pupils, rather than having to learn by own mistakes too many times on things that don’t work...”

Marit thinks that when it comes to seminars, PE is not as prioritised as other subjects:

“I think things are well organised. I would like to have some more courses though... A bit more seminars in the PE subject... That it could have been more of. Because it’s a lot of that in core subjects like Norwegian, mathematics and English. So the PE subject should get some more focus I think.”

The literature review in this thesis has not dealt with teachers’ access to online resources and I have not encountered research about this in my search for literature. It has still emerged as challenge during the research interviews. One challenge that has emerged, is the lack of time to attend seminars as Bente describes, which is in line with Hansen’s findings (2017, p. 36), where teachers were hindered to have seminars and meetings on assessment due to time barriers. Marit also adds that she thinks seminars for PE is not prioritised when comparing with other subjects. If PE is not prioritised and less resources are spent on PE and the implementation of AFL in the subject, this clearly represent a challenge considering the discrepancy between current practices and LOA. There seems to be a great need for further AFL seminars, as highlighted in the report about the Norwegian AFL project (Utdanningsdirektoratet, 2018). Both Bjørg and Tormod miss more online resources and they are not happy with the resources provided by UDIR. They are the two youngest and least experienced teachers in the study, which are both eager to learn more, but find it hard to locate resources about assessment. Tormod questions why there is no textbook for the subject. It is evident that teachers do not feel they learn enough about assessment in teacher education, they do not feel there are enough seminars, or that they have time to attend them. In addition to this, they feel there is not enough online resources. These are challenges that questions the reinforcement of AFL practices in Norwegian PE classes.

4.3.9 Class size and time to assess

Another challenge to LOA that emerged from the interviews was class sizes. Big classes limit the time teachers have to see and instruct every pupil.

Sonja finds the feedback coming into conflict with instructing when classes are too big:

“we have 60-minute sessions... Then the pupils are active in maybe... max 40 minutes. We have big classes now, 29-30 pupils. You can't assess each pupil every session. No chance. Then they won't learn anything, that's the issue... You can't give feedback to everyone, I'm not able to that while simultaneously instructing...”

She clarifies more about not having enough time:

“Well, I see them too little... We teachers can't complain that we don't have enough time... (The issue is) you see the pupils too little in class, you don't see them enough to assess them. You can't take them out of other classes to give them feedback about how they were during PE...”

Stine says the time it takes to assess the pupils eats of the time she should have used for good enough feedback during sessions:

“You constantly need to focus on how they perform at all times and I think the pupils also feel they are constantly assessed... So getting enough time for... and good enough quality... In a way you need to take time away from your instructing to give good enough feedback...”

Bente, who provides individual feedback after each session, describes to process:

“we have PE for two school hours, one hour and 45 minutes and I write down feedback in two or three hours after the sessions... I write more comprehensive feedback when I have (just seen them) ... So that takes a lot of time, but at the same time, I use less time during the sessions, cause you're supposed to teach as well.”

Marit has smaller classes, but have experience with bigger classes previously:

“small classes... 17, 16... But the last school I worked at I had 60 pupils in total, who was my responsibility, and that way too many. It just... You can't even learn all the

names, and that I find unsustainable.”

Tora finds it hard to see everyone in the class:

“now I have a class where there are 28 pupils... That’s a lot. So it’s clear, you see... maybe you see the extremes, right. Those that excels or those that maybe don’t do too much, but those in the middle kind of... get invisible.”

Johann feels a pressure on his own time:

“I assess them in my spare time. That’s when I have time to assess. Certainly as a contact teacher, because the available space on my schedule goes to my class anyway, on pupil conversations... The basis for assessment is not enough, it’s incredibly short time...”

In order to provide meaningful feedback, I assume that a prerequisite for this is to have the opportunity to see all pupils. The responses from the teachers suggest some issues regarding this. Thorleif, Hallgeir, Stine and Tormod highlights that the time it takes to assess everyone goes at the expense of learning activities and thereby creates a big challenge to LOA. Tormod describes issues with providing quality feedback for everyone in a big class, which is highlighted by Jonskås’s (2009) and Hansen’s (2017) qualitative interviews with teachers and Mørken’s quantitative questionnaire (2010). Both Marit, Tora and Johann discuss the difficulty in seeing all pupils when classes get too big. Jonskås (2009, p. 55) showed evidence of a teacher who was responsible for the assessment of over 300 PE pupils. Internationally, we have seen pupil to teacher ratios can get up to 400 pupils per teacher (Turner et al, 2017), which makes individualised feedback and instructions very challenging. As we have seen in previous results section and which we will see in the coming section on private schools, one teacher who seems to have little trouble with feedback is also the teacher who sees the pupils the most, in the sports school. It is obvious that the third formative strategy by William and Thompson will be much easier to implement in smaller classes, which is also highlighted by Marit, who suggests things are better in her current school, with smaller classes.

The boundaries related to the timetable has been identified by Kirk (2010, p. 104) and is also evident in this thesis. Sonja explains, the issue is not necessarily that teachers do not have enough time to assess, but that they do not get the chance to see pupils enough to provide meaningful feedback and assessment. However, many teachers, such as Thorleif and Johann

also feel the time pressure on their own time, which makes me question whether all PE teachers sit down on their spare time to provide feedback to pupils. If the structures surrounding PE does not allow enough time for many teachers to give feedback, there is no guarantee that all teachers have the resources or the capacity to provide quality feedback to pupils. Thorleif explains that they need to find methods, which get pupils to understand what they are working towards and what they are supposed to do in the subject, which showcases a great potential for formative strategies to be more strongly implemented.

4.3.10 Facilities and resources

Teachers were also asked about the facilities and geographical surroundings, which emerged as a topic where many felt limited opportunities within PE.

Thorleif talks about the difficulty when it comes to outdoor recreation:

“Outdoor recreation is tough... It would be nice if they could walk 200 m in flat terrain to get to a forest instead of 1km uphill which is the case for us... That’s a challenge we have, so outdoor recreation is not prioritised, so we get a limited basis for assessment in that.”

Arnfinn from the private school feels he is in a privileged position:

“we are lucky. We have a lot of equipment, which gives us the opportunity to vary our activities quite a lot, but I know from my previous job, there wasn’t a lot of equipment, so our activities were fewer and less diverse...That makes the subject more boring to many, because if you don’t find football fun...Giving the pupils equal opportunity to express themselves despite having different backgrounds, that’s linked to economics, facilities, resources.”

Tormod feels frustration about the facilities and resources with implications for his opportunity to provide feedback:

“I feel the resources makes it easier for some schools to assess better than for example in this school, where the facilities are not that great... To get a better assessment situation the facilities are very important and it’s not well facilitated for all schools. It’s about fair distribution of the budget that could have been different. In arts and craftsmanship for example, the pupils are divided into two groups in their

classes. Much higher teacher density and you have the opportunity to provide better feedback so it's easier for the pupils to show their best."

Thorleif takes up the issue geographical barriers may possess to activities in the competence goals such as outdoor recreation. Although not strictly related to assessment, it is obvious that LOA in relation to the learning goals in outdoor recreation will be difficult or impossible if facilities do not allow the activity to be conducted. Also, literature have shown that the lack of such activities may overemphasise sports in PE, with negative effects for some groups of pupils (Jonskås, 2009; Omholt, 2015). Although Arnfinn feels he has good opportunities in his current job, he describes his previous job in public school with limited opportunities to design a learning environment that gave all pupils the chance to express themselves. For Bente, the frustration is linked to space. Tormod compares PE to arts and craftmanship, which he suggests have two teachers per class, which creates a much more opportunity to provide feedback to each pupil. This is linked to the same challenge as class sizes. Wiken (2011, p. 61) showed the lack of prioritisation of activities such as orienteering, which obviously reduces the learning potential in such areas.

4.3.11 The biggest challenge

Finally, in the research interviews, teachers were asked what they saw as the biggest challenge with assessment in PE, as a way of understanding what issue stresses them the most. Two teachers have been left out of this section, as their responses were not linked to the research question. The results are presented below in table-format.

Teacher	Biggest issue	Sample extract
Thorleif	Assessment in outdoor recreation	<i>"the assessment in for example outdoor recreation is a bit thin... Because they want us to assess it, but I find that very abruptly and ill-founded."</i>
Sonja	Lack of time	<i>"that's the time aspect. Time aspect and time aspect. That you... you see them too little in a way. They have two hours of PE a week, there's no more than that. And then you're supposed to grade them based on that, right."</i>

Arnfinn	Lack of time	<i>“the time aspect. Obviously. I think that you can’t try getting done too much... Uhm, because the risk of failing is quite high. That you can’t do it (the assessment) good enough for everyone. And that it gets unfair...”</i>
Bjørg	Level differences among pupils	<i>“I find it hard when there are so many levels of ability in a class. That is a challenge. For there are many that don’t feel safe, if you’re not athletic for example. And those that are athletic don’t get so much out of it always, because you have to facilitate for those that aren’t athletic as well.”</i>
Stine	Lack of time	<i>“it’s getting time for it. Time enough... Because it’s non-existing really. In a way, you have to eat of the time that you have for the teaching to give a good enough feedback.”</i>
Bente	Space, equipment, lack of time	<i>“for us, maybe space and equipment. Space particularly. Time for swimming. We have... we have swimming but very little time... Yeah, space, equipment and time.”</i>
Tormod	Lack of time	<i>“uhm that is to see the pupils enough to provide them with a fair assessment and follow up on them enough. Especially those that don’t come and ask, so that they all the time know where they’re at and what they can work on. And to come and talk with me about things that are hard...”</i>
Marit	Time for cooperation between teachers	<i>“I don’t think there are that many (issues), but it would be nice if we had been more teachers that could cooperate more, if we had time for that. The time perspective on that, I think.”</i>
Tora	Class sizes	<i>“the biggest challenge right now I feel is having such a big class... To have 20 versus 28, that’s a big difference... To have eight extra pupils in PE for example, that makes... it limits the opportunity to what you can do...”</i>

Johann	Unmotivated pupils	<i>“what I struggle the most with is the pupils that sit on the sideline... Getting those that don't want to attend, getting in those that are afraid... and who refuses... That's very hard...”</i>
--------	--------------------	--

This final section of the results and analysis confirms that the time aspect seems to represent the biggest challenge for LOA from teachers' perspectives. While this is not a quantitative study and one must be careful to generalise, it regardless stands out as a major barrier to implementing LOA in PE. Linked to this is also class sizes and the way facilities allow teachers to see all pupils. If the space is too small for a big class, it gets hard for teachers to see everyone express themselves in the subject. There are also other aspects that stand out for the teachers, such as level differences among pupils. As we have seen in the literature review, PE pupils are diverse, and the subject has traditionally been designed in a way that has excluded or at least made many pupils dislike the subject. The second formative strategy, which has not been implemented by UDIR in the AFL project (Utdanningsdirektoratet, 2018), calls for tasks that elicit evidence of pupils' learning. This must be the case for all groups of pupils, not only for active pupils that enjoys competing.

4.4 Private vs public schools

This study included two private schools which adds in one comparative element of potential differences between teachers' responses from private and public schools. Due to the small sample it is impossible to generalise on this, but any differences may add to knowledge on the topic and multiple perspectives are important for a broader understanding. Teachers from private schools must work by the same goals, rules and regulation as other schools. The question is whether they have more resources or better opportunities to facilitate for LOA, considering they take tuition fees from pupils.

Arnfinn, from the elite sports school, talks about the large amount of time they have in PE in his school and the time they have for practice. He suggests they have much better opportunities to see the pupil, because they are practising the subject every day:

“We have so much more time to practice than in a public school... We practice every day, so the room for improvement is quite big. At the same time, we have, a bigger scope in content, we try out more activities than perhaps public schools are able to.”

It is clear from the interview that Arnfinn is in a privileged situation where his school has many more hours of physical activity than public schools. He also states that he thinks they can do more diverse types of activities and make the subject more interesting this way. Arnfinn talks about the benefits of having motivated pupils and the opportunities this creates for feedback and learning:

“Most of them finds PE fun and wants to train every day, wants to get better. And then you are much more responsive for guidance”.

This highlights one of the big differences between a private sports school and a standard public school, where the private school has the privilege of having largely motivated pupils, who desire your feedback. In terms of the opportunity to implement LOA, Arnfinn seems to be the teacher who enjoys the best potential for LOA in his school. This is due to the large amount of time he can focus on practice and learning, as opposed to other teachers, which takes pressure away from the need for documentation, tests and observation. As quoted in the section about good practice, he can facilitate for individual pupils so that everyone is able to perform a cartwheel in gymnastics and he has time to give everyone feedback. The teachers on public schools, on the other hand, seem to be unable to do this due to restricted time. Therefore, I argue that having time for LOA in PE is one of the biggest challenges for implementing it. However, Arnfinn also enjoys the benefit of having eager pupils that are highly active and he does not have to spend time on pupils that struggle with amotivation (Ryan & Deci, 2020). Also, it is not realistic that time for PE in the learning plan get drastically higher any time soon, hence improving LOA in PE need to address more than increasing the amount of time.

Björg, who works in an academic private school, says that there are pros and cons with working in her private school, as she is the only PE teacher at her school, which allows here much freedom in how to design the PE classes:

“I have the opportunity to take them wherever I want, even if it perhaps costs some money. The climbing centre or hiring a sports centre where there are some extra facilities, or we can hire a professional coach in swimming for example.”

Björg speaks of the financial resources they have compared with public schools, where they can hire facilities or coaches in areas where the teacher lacks competence, which creates a lot of flexibility and potential to learn activities that are out of reach for public schools. She

brings up the opportunities associated with being a small school, which will also be relevant for small public schools. She states that they are 40 pupils and each PE class is only 20 pupils, where public schools typically have 30, according to her. She talks about the opportunity she has to talk to each pupil:

“I can just go and talk to them whenever, if they sit and work and have some study time or something. I don’t have to talk to them in the PE session and that’s a good thing, because we don’t have that many hours of PE a week”.

Bjørg touches into the opportunity of spending time that is not devoted to activity to provide feedback and guidance for the subject. This gives her an opportunity to focus more on activity and learning during the classes. While Bjørg talks about the flexibility to conduct diverse activities due to having more resources, there seem to be few factors that separate Bjørg’s context from that of a public school. In this sense, the difference between private and public schools seem to depend on the type of private school. While Arnfinn at the sports school have much time to facilitate for learning through LOA at his school, Bjørg have more similar responses to teachers from public schools. She has more resources available, but this does not alleviate the challenges such as restricted time for the subject, documentation needs, PETE and unmotivated pupils.

5 Conclusion

The intention of the thesis has been to explore the assessment practices among secondary education PE teachers and the challenges to implementing LOA. The research questions have been responded to through semi-structured qualitative research interviews with twelve Norwegian PE teachers from ten schools. The practices and challenges have been diverse and the theoretical framework on formative assessment by Black and William (2009) and the formative strategies of William and Thompson (2007) have been used both to filter and analyse the responses from the interviews. Below I will summarise the findings for each of the research questions.

5.1 Research question 1: what forms of assessment practices do PE teachers use, and to what extent do they align with learning-oriented assessment practices?

The evidence about the first formative strategy about sharing learning criteria with pupils is limited, as not all interviews contain information about this. However, the teachers that did discuss this highlight variance to the degree of its implementation. Teachers seem to recognise the importance of this strategy, but some sees this as an area to improve. The second formative strategy in this thesis was linked to assessment methods, such as testing and authentic assessment, as the methods determine the degree to which pupils can showcase their learning and abilities. In the interviews, it became clear that formative assessment or AFL are not concepts that are widely used in PE. Lopez-Pastor et al (2012) uses the term ‘alternative assessment’ as opposed to the traditional way (testing, observation, etc) to bring forward innovative assessment methods that focus more on learning in PE. In my thesis, it has been hard to identify ‘alternative’ approaches, however, I have shown evidence of teacher practices in line with formative strategies (William & Thompson, 2007). These methods include rich tasks, where pupils have needed to use multiple skills to solve a task, and performance assessment, where the teacher lets the pupil practice and takes note of the best performance without informing the pupil about the assessment situation. Modifying the rules of the game and attempts to level the playing field in PE are also evidence of practices in line with the second formative strategy. However, it has been clear from the literature that such evidence is not common for all teachers.

While research suggest testing is widely used in PE, most teachers in this study avoid testing, which indicates that assessment practice in PE is gradually reforming. One teacher highlighted the use of skill tracks, which is described as a testing situation. Several teachers

use tests formatively, by giving pupils the same test twice during the school year. Much of the assessment methods in the findings constitute authentic assessment, where the use of observation and video evaluates realistic situations such as a game in teams sports, as opposed to constructed testing situations. Observation seems to represent the most widely used assessment method among the teachers. One key finding is that notetaking, or documenting is a big part of the teachers' assessment and they devote much time in documenting pupils' performances.

The third formative strategy is about feedback and most teachers manage to provide formative assessment to pupils. However, the degree to which they are able to do this vary. The teacher from the sports school have much time to give feedback and another teacher spends much time providing written feedback following each session, while other teachers tries their best to provide feedback with a varying degree of success. From the section about challenges, it is clear there are structural factors limit the degree to which feedback is given to pupils. The fourth formative strategy is to using peers as instructional resources for one another, where evidence clearly show that this method receives little attention among teachers. A few teachers provide evidence of this, particularly within dance and gymnastics and one teacher allows pupils to take control over sessions occasionally. Peer assessment was not used by teachers, and many were sceptical towards the use of it, in line with literature on the subject. The fifth formative strategy deals with pupils' ownership of their own learning, which is linked to self-assessment. While there is much reference to this in the interviews, this appears to be, as suggested by research, a matter of seeing how pupils' grading of themselves align with the teacher's grading. Nor is self-assessment used regularly, but rather seems to be a check point that teachers need to provide at some point during the school year. There are no indications from the teachers' responses that pupils are owners of their own learning in contemporary PE.

The evidence in the first research question show, in line with literature explored in this thesis, that there is a long way to go for LOA assessment methods in PE. There appears to be great variance to the degree pupils know what they are supposed to learn, and assessment methods seem to be more about assigning grades than about facilitating learning. Pupils are only occasionally engaged in reflecting about their own learning. However, there are evidence of good practice and testing for grading purposes are not apparent in the teachers' responses.

5.2 Research question 2: are there barriers to teachers using learning-oriented assessment practices in their PE practice, and what kinds of support or advice do physical education teachers need to develop educationally sound, successful and sustainable forms of assessment?

Challenges to the implementation of LOA is an area without great coverage in the literature. While the literature review highlights many challenges for assessment and LOA respectively, no literature I have encountered goes into this specifically. The results from the interviews show that challenges to LOA are many and diverse, some of which represents real barriers to the implementation of LOA. One of the clearest indicators to the lack of LOA in PE is the teachers' responses to questions about whether they think learning is happening in PE and whether they think grades in the subject reflect a learning curve, or merely the status quo level of the pupil. Most of the teachers are sceptical regarding the presence of learning in the subject. Many teachers suggest a conflicting interest between assessment and learning in PE, and some think the assessment hinders learning among pupils, as it takes away time that should be spent facilitating learning. A couple of teachers say there is not enough time for pupils to achieve learning goals in the subject and only a single teacher responds solely positively to the question on learning.

The interviews show how pupils and their parents challenge teachers and their assessment, which together official regulations (Lovdata, n.d.) create a strong need for documentation of pupils' performance and abilities in the subject. This, I argue, is one of the biggest challenges for LOA in PE. Based on the teachers' responses, the need to document pupils seems to pressure teachers to facilitating activities that focuses on performance, such as skill tracks and activities where the teacher takes a background position to document pupils' performances. This take focus away from mastery and practice, eats time they have for instruction and feedback and takes a lot of time for teachers. The results support the literature in that assessment is stressful for many pupils, seen from the teachers' perspectives. In addition to this, teachers imply that the extrinsic motivation of grades is the most important to many pupils, which undermines intrinsic motivation for the subject (Ryan & Deci, 2020). Also supported by literature is the lack of participation of some groups of pupils. While this clearly represent a barrier to LOA of these pupils, it also requires time and attention of teachers', who uses much time trying to get these unmotivated pupils to attend, taking away time they have for feedback and facilitating LOA.

While some teachers enjoy the decentralised system and the flexibility associated with the competence goals, some find the goals and guidelines too vague. One teacher also highlights a challenge where the local learning goals plan highlights technical aspects of the subject, which facilitates for PE-as-sport-techniques (Kirk, 2010). Most teachers agreed that assessment training in their teacher educations was inadequate and too theoretical, which created issues for some once they started working in the field. The lack of practical training may create a challenge for many teachers, balancing the needs for documentation, teaching and providing meaningful feedback for pupils. In addition to this, some teachers find the online resources for teachers by UDIR to be too few and not helpful enough with regards to assessment. Two major and interconnected areas that comes up as major challenges in the interviews are class sizes and time to assess. All teachers but one finds their classes too big to be able to see all pupils, facilitate learning for these and provide regular, good quality feedback. Big classes and huge amounts of documentation take up much of the teachers' time, both of their personal time, but more importantly, with regards to LOA, it steals from the time they have to instruct and give feedback to pupils during the sessions. Due to the teachers having other subjects, they do not have enough time to follow up each pupil, with negative implications for learning. Alongside class sizes, there are other physical constraints that challenge the LOA in PE, such as the lack of areas to learn outdoor recreation or swimming, the lack of equipment which can limit PE sessions to ball sports, and the lack of space, which creates challenges for the teacher to 'see' all pupils.

From the question about the biggest challenge and based on my interpretations of the interviews all together, the biggest barrier to LOA in Norwegian PE seem to be the limited time they have to see and instruct their pupils. This lack of time seems to be caused by the pressure teachers experience to document pupils' performances, short and few sessions in the subject and big class sizes. There are also other challenges such as inadequate training and education, lack of facilities and pupils' lack of motivation. The teachers' responses reveal that there are great concerns when it comes to the concept of learning in PE and this thesis highlights the need to look critically at the time and resources teachers have available to design educationally sound and sustainable methods of assessment.

5.3 Private vs public schools

While the comparative element between teachers from private and public schools in this thesis is not generalisable, due to a limited sample, a few remarks have emerged from the

teacher interviews. The responses from the teacher in the private sport school suggest that they had much more time for PE and thereby much more time for practice and feedback. The teacher's statements showed promising results in relation to LOA compared with other responses from other participants. This finding has helped support the argument that the time aspect represents a major barrier to the implementation of LOA. Responses from this teacher and from the teacher from the academic private school also show that they have more resources to spend on the subject. However, responses from the academic private school teacher show that many of the challenges are the same as public school teachers face. This has brought me to the conclusion that the difference between private and public secondary, when it comes to assessment practices and LOA challenges, depend on the type of private school and the amount of time their teachers can spend on PE.

5.4 Contribution to the field

By exploring the two research questions through interviewing PE teachers, this thesis has shed light to some research areas there are limited research about. Much of the literature explored focus on just and fair assessment, the inability of PE to reform its' assessment practices, detrimental effects of assessment and evidence on the promising results of LOA. Little research focus on the assessment practices used in PE, except from the use of testing. This thesis adds to the knowledge about this area. By comparing the teachers' responses to the theoretical framework on formative assessment, this thesis has illuminated the degree to which assessment practices in Norwegian secondary education PE is linked to formative assessment. When it comes to the second research question, the results from this thesis helps reducing the knowledge gap about challenges to implement LOA in PE. While research confirms promising results of formative assessment and much research highlights problematic sides of assessment in PE, this thesis contributes to knowledge about why PE teachers struggle with implementing LOA.

5.5 Future research

There are many aspects of the findings in this thesis that require more research. To do a more detailed and thorough evaluation of the teachers' assessment practices, this require spending more time on this in interviews and preferably observation of PE classes. All formative strategies by William and Thompson should be explored more for PE in Norway and elsewhere. Teachers seem to lack a methodology on how to create opportunities for all PE pupils to show their abilities, as well as methods to make PE more inclusive. I would like to

see more case study research on how teachers can facilitate this in the subject, which would reduce some of the challenges linked to participation and motivation in this thesis. More research should go in depth about feedback in Norwegian PE classes. The nature of feedback teachers give could itself be a master thesis. How often do pupils receive feedback? What type of feedback is it? How well does it support learning? Based on the results of this thesis, it seems like teachers lack methodology or guidelines on how to do self-assessment, where many teachers treat self-assessment as a reality-orientation tool for pupils, rather than a tool to engage the pupils in deeper learning. I suggest more concrete research on self-assessment in PE, particularly quantitative studies on the frequency of its use. With regards to teacher education, the results from this thesis imply that the training teachers get on assessment is inadequate. Future research should look more into this, such as, to what degree do students learn about AFL practices and the practical implementation of these? I also suggest more research on the needs of teachers in terms of learning material and online resources, which could provide guidelines on what teachers require to improve LOA. Finally, I would like to see bigger studies on the major barriers to LOA, identified in this thesis, the lack of time to for LOA, as well as the pressure teachers face to document pupils' performances. What actions are required for policymakers to improve the potential for LOA?

Reference List

- Andrews, T., & Johansen, V. (2005). "Gym er det faget jeg hater mest". *Norsk pedagogisk tidsskrift* 89(04), 302–314. Retrieved from https://www-idunn-no.ezproxy.uio.no/npt/2005/04/gym_er_det_faget_jeghatermest
- Annerstedt, C. (2008). "Physical Education in Scandinavia with a Focus on Sweden: A Comparative Perspective." *Physical Education and Sport Pedagogy* 13(4), 303–318. doi: [10.1080/17408980802353347](https://doi.org/10.1080/17408980802353347)
- Annerstedt, C., & Larsson, S. (2010). "I Have My Own Picture of What the Demands Are ...": Grading in Swedish PEH — Problems of Validity, Comparability and Fairness." *European Physical Education Review* 16(2), 97-115. doi: [10.1177/1356336X10381299](https://doi.org/10.1177/1356336X10381299)
- Arnesen, T. E., Nilsen, A. K., & Leirhaug, P. E. (2013). "Den læreplanen som ikkje kan tilpassast mi undervisning, finst ikkje.": vurdering og undervisning i kroppsøving etter kunnskapsløftet." *Tidsskriftet FoU i praksis* 7. 9-32. Retrived from <http://hdl.handle.net/11250/225309>
- Augestad, P. (2003). "Architecture and the Education of the Body: The Gymnasium in Norwegian Physical Training, 1889-1930." *The International Journal of the History of Sport* 20(3). 58–76. doi: [10.1080/09523360412331305783](https://doi.org/10.1080/09523360412331305783)
- Bakken, A. (2019). *Idrettens posisjon i ungdomstida. Hvem deltar og hvem slutter i ungdomsidretten?* (NOVA Report 2/19). Retrieved from <http://hdl.handle.net/20.500.12199/1298>
- Berry, R. (2011). Assessment Reforms Around the World. In R. Berry & B. Adamson (Eds.), *Assessment Reform in Education: Policy and Practice, Education in the Asia-Pacific Region: Issues, Concerns and Prospects*, (pp. 89-102). Dordrecht: Springer Netherlands. doi: [10.1007/978-94-007-0729-0_7](https://doi.org/10.1007/978-94-007-0729-0_7)
- Bezeau, D., Turcotte, S., Beaudoin, S., & Grenier, J. (2020). "Health Education Assessment Practices Used by Physical Education and Health Teachers in a Collaborative Action

- Research.” *Physical Education and Sport Pedagogy* 25(4). 379–393. doi:
[10.1080/17408989.2020.1725457](https://doi.org/10.1080/17408989.2020.1725457)
- Birch, J. E. (2016). Subjektiv vurdering er rettferdig vurdering. In E. E. Vinje (Ed.), *Kroppsøvingdidaktiske utfordringer*, (pp. 51-64). Tønsberg: Cappelen Damm Akademisk.
- Black, P., & Wiliam, D. (1998). “Assessment and Classroom Learning.” *Assessment in Education: Principles, Policy & Practice* 5(1). 7–74. doi:
[10.1080/0969595980050102](https://doi.org/10.1080/0969595980050102)
- Black, P., & Wiliam, D. (2009). “Developing the Theory of Formative Assessment.” *Educational Assessment, Evaluation and Accountability* 21(1). 5–31. doi:
[10.1007/s11092-008-9068-5](https://doi.org/10.1007/s11092-008-9068-5)
- Bryman, A. (2016). *Social Research Methods* (5th ed.). Oxford: University Press.
- Butler, S. A., & Hodge, S. R. (2001). “Enhancing Student Trust Through Peer Assessment in Physical Education.” *Physical Educator* 58(1). 30-42. Urbana: Sagamore Publishing LLC.
- Cambridge Assessment. (n.d.). *Learning-Oriented Assessment*. Retrieved October 26, 2020, from <https://www.cambridgeenglish.org/research-and-validation/fitness-for-purpose/loa/>
- Cassady, H., Clarke, G., & Latham, A. M. (2004). “Experiencing Evaluation: A Case Study of Girls’ Dance.” *Physical Education and Sport Pedagogy* 9(1). 23–36. doi:
[10.1080/1740898042000208106](https://doi.org/10.1080/1740898042000208106)
- Chen, W. (2005). “Examination of Curricula, Teaching Practices, and Assessment through National Standards.” *Physical Education and Sport Pedagogy* 10(2). 159–80. doi:
[10.1080/17408980500105056](https://doi.org/10.1080/17408980500105056)
- Engvik, G. (2010). Elevvurdering i Faget Kroppsøving - Noen utfordringer. In S. Dobson & R. Engh (Eds.), *Vurdering for læring i fag*, (pp. 201–213). Kristiansand: Høyskoleforlaget.
- Gardner, John. 2012. *Assessment and Learning* (2nd ed.). London: Sage Publications.

- Georgakis, S., & Wilson, R. (2012). "Australian Physical Education and School Sport: An Exploration into Contemporary Assessment." *Asian Journal of Exercise & Sports Science* 9(1). 37–52.
- Græsholt, S. A. (2011). *Elevvurdering og vurderingskultur i kroppsøving: hvordan erfarer elever i videregående skole elevvurdering, og hvilke vurderingskulturer kommer til uttrykk i kroppsøvingfaget?* (Master's thesis, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/171538>
- Guest, G., Bunce, A., & Johnson, L. (2006). "How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability." *Field Methods* 18(1). 59–82. doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)
- Hansen, E., Å. (2017). *Vurdering i kroppsøving – en stor utfordring! : Hvordan erfarer lærere og elever i ungdomsskolen elevvurdering, og hvilke vurderingskulturer ligger til grunn i vurderingen?* (Master's thesis, University of Agder, Norway). Retrieved from <http://hdl.handle.net/11250/2454511>
- Harlen, W. (2012). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 171-184). London: SAGE Publications. doi: [10.4135/9781446250808.n11](https://doi.org/10.4135/9781446250808.n11)
- Hay, P. J., & Macdonald, D. (2010). "Evidence for the Social Construction of Ability in Physical Education." *Sport, Education and Society* 15(1). 1–18. doi: [10.1080/13573320903217075](https://doi.org/10.1080/13573320903217075)
- Hondrich, A. L., Decristan J., Hertel S., & Klieme E. (2018). "Formative Assessment and Intrinsic Motivation: The Mediating Role of Perceived Competence." *Zeitschrift Für Erziehungswissenschaft* 21(4). 717–34. doi: [10.1007/s11618-018-0833-z](https://doi.org/10.1007/s11618-018-0833-z)
- Hopfenbeck, T.N., Tolo A., Florez T., & El Masri, Y. (2013). *Balancing Trust and Accountability? The Assessment for Learning Programme in Norway* (OECD Education Working Papers 97). Retrieved from https://read.oecd-ilibrary.org/education/balancing-trust-and-accountability-the-assessment-for-learning-programme-in-norway_5k3txnpqlsnn-en#page1

- Hortigüela Alcalá, D., & Hernando Garijo, A. (2017). "Teaching Games for Understanding: A Comprehensive Approach to Promote Student's Motivation in Physical Education." *Journal of Human Kinetics* 59. 17–27. doi: [10.1515/hukin-2017-0144](https://doi.org/10.1515/hukin-2017-0144)
- Ingebrigtsen, J. E., & Mehus, I. (2006). Kroppsøving og idrett for alle: hvordan nå målet?. In H. Sigmundsson & J. E. Ingebrigtsen (Eds.), *Idrettspedagogikk* (pp. 33–46). Oslo: Universitetsforlaget.
- Jaakkola, T., Yli-Piipari, S., Watt, A., & Liukkonen, J. (2016). "Perceived Physical Competence towards Physical Activity, and Motivation and Enjoyment in Physical Education as Longitudinal Predictors of Adolescents' Self-Reported Physical Activity." *Journal of Science and Medicine in Sport* 19(9). 750–54. doi: [10.1016/j.jsams.2015.11.003](https://doi.org/10.1016/j.jsams.2015.11.003)
- Jaakkola, T., John Wang, C. K., Soini, M., & Liukkonen, J. (2015). "Students' Perceptions of Motivational Climate and Enjoyment in Finnish Physical Education: A Latent Profile Analysis." *Journal of Sports Science & Medicine* 14(3). 477–83. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4541109/>
- James, M. & Lewis, J. (2012). Assessment in Harmony with our Understanding of Learning: Problems and Possibilities. In J. Gardner (Ed.), *Assessment and Learning* (pp. 187–205). London: SAGE Publications.
- Jonskås, K. (2009). Elevvurdering i kroppsøving: hvordan tolker og praktiserer lærere i den videregående skole elevvurdering etter innføring av Kunnskapsløftet? (Master's Thesis, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/171420>
- Kirk, D. (2010). *Physical Education Futures*. London: Routledge.
- Kolbeinsen, L. S. (2016). 'Det viktigste er jo hvordan du presterer i gymmen': en kvalitativ studie av elevers erfaringer med kroppsøvingfaget. (Master's Thesis, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/2400417>

- Lagestad, P. A. (2017). "Er gutter bedre enn jenter i kroppsøving? -En studie av jenter og gutters kroppsøvingsskarakterer i den videregående skolen." *Acta Didactica Norge* 11(1). 1-21. doi: [10.5617/adno.2609](https://doi.org/10.5617/adno.2609)
- Lagestad, P. A., & Sørensen, A. (2018). "Longitudinal Changes in Sports Enjoyment among Adolescents." *Journal of Physical Education and Sport* 18(1). 89-97. doi: [10.7752/jpes.2018.01011](https://doi.org/10.7752/jpes.2018.01011)
- Leirhaug, P. E. (2016). 'Karakteren i seg selv gir jo ikke noe læring': en empirisk studie av vurdering for læring i kroppsøving ved seks videregående skoler i Norge (Doctoral Dissertation, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/2419782>
- Leirhaug, P. E., & Annerstedt, C. (2016). "Assessing with New Eyes? Assessment for Learning in Norwegian Physical Education." *Physical Education and Sport Pedagogy* 21(6). 616–31. doi: [10.1080/17408989.2015.1095871](https://doi.org/10.1080/17408989.2015.1095871)
- Leirhaug, P. E., & MacPhail, A. (2015). "'It's the Other Assessment That Is the Key': Three Norwegian Physical Education Teachers' Engagement (or Not) with Assessment for Learning." *Sport, Education and Society* 20(5). 624–40. doi: [10.1080/13573322.2014.975113](https://doi.org/10.1080/13573322.2014.975113)
- López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2012). "Alternative Assessment in Physical Education: A Review of International Literature." *Sport, Education and Society* 18(1). 57–76. doi: [10.1080/13573322.2012.713860](https://doi.org/10.1080/13573322.2012.713860)
- Lovdata. (n.d). *Forskrift Til Opplæringslova - Kapittel 3. Individuell Vurdering i Grunnskolen Og i Vidaregåande Opplæring - Lovdata*. Retrieved November 1, 2020 from https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_5#KAPITTEL_5
- Macdonald, D. (2011). "Like a Fish in Water: Physical Education Policy and Practice in the Era of Neoliberal Globalization." *Quest* 63(1). 36–45. doi: [10.1080/00336297.2011.10483661](https://doi.org/10.1080/00336297.2011.10483661)

- Macphail, A., & Halbert, J. (2010). "We Had to Do Intelligent Thinking during Recent PE": Students' and Teachers' Experiences of Assessment for Learning in Post-Primary Physical Education." *Assessment in Education* 17. 23–39. doi: [10.1080/09695940903565412](https://doi.org/10.1080/09695940903565412)
- Mintah, J. (2003). "Authentic Assessment in Physical Education: Prevalence of Use and Perceived Impact on Students' Self-Concept, Motivation, and Skill Achievement." *Measurement in Physical Education and Exercise Science* 7. 161–74. doi: [10.1207/S15327841MPEE0703_03](https://doi.org/10.1207/S15327841MPEE0703_03)
- Moen, K. M. (2011). *'Shaking or Stirring'? A Case-Study of Physical Education Teacher Education in Norway* (Doctoral Dissertation, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/171359>
- Møller, J., & Skedsmo, G. (2013). "Modernising Education: New Public Management Reform in the Norwegian Education System." *Journal of Educational Administration and History* 45(4). 336–53. doi: [10.1080/00220620.2013.822353](https://doi.org/10.1080/00220620.2013.822353)
- Mørken, B. (2010). *Assessment in Physical Education* (Master's Thesis, Telemark University College, Norway). Retrieved from <http://hdl.handle.net/11250/2438747>
- Nesheim, H. (2008). *Selvbestemmelse og autonomistøtte i videregående skole: en studie om motivasjon for kroppsøving i videregående skole* (Master's Thesis, Nord-Trøndelag University College, Norway). Retrieved from <http://hdl.handle.net/11250/146962>
- Omholt, O. K. J. A. (2015). *Motivasjon i kroppsøvingfaget hos jenter ved videregående skole* (Master's Thesis, University of Stavanger, Norway). Retrieved from <http://hdl.handle.net/11250/299355>
- Ommundsen, Y., & Eikanger S. K. (2007). "Autonomy–Mastery, Supportive or Performance Focused? Different Teacher Behaviours and Pupils' Outcomes in Physical Education." *Scandinavian Journal of Educational Research* 51(4). 385–413. doi: [10.1080/00313830701485551](https://doi.org/10.1080/00313830701485551)
- Perrenoud, P. (1998). "From Formative Evaluation to a Controlled Regulation of Learning Processes. Towards a Wider Conceptual Field." *Assessment in Education: Principles, Policy & Practice* 5(1). 85–102. doi: [10.1080/0969595980050105](https://doi.org/10.1080/0969595980050105)

- Prøitz, T. S., & Borgen, J. (2010). *Rettferdig standpunktvurdering - det (u)muliges kunst?: læreres setting av standpunktkarakter i fem fag i grunnopplæringen* (NIFU Step Report 16/2010). Retrieved from <http://hdl.handle.net/11250/279131>
- Ramprasad, A. (1983). "On the Definition of Feedback." *Behavioral Science* 28. 4–13. doi: [10.1002/bs.3830280103](https://doi.org/10.1002/bs.3830280103)
- Ryan, R. M., & Deci, E. L. (2020). "Intrinsic and Extrinsic Motivation from a Self-Determination Theory Perspective: Definitions, Theory, Practices, and Future Directions." *Contemporary Educational Psychology* 61. doi: [10.1016/j.cedpsych.2020.101860](https://doi.org/10.1016/j.cedpsych.2020.101860)
- Säfvenbom, R., Haugen, T., & Bulie, M. (2015). "Attitudes toward and Motivation for PE. Who Collects the Benefits of the Subject?" *Physical Education and Sport Pedagogy* 20(6). 629–46. doi: [10.1080/17408989.2014.892063](https://doi.org/10.1080/17408989.2014.892063)
- Sandvik, L. V., Engvik, G., Fjørtoft, H., Langseth, I. D., Aaslid, B. E., Mordal, S., & Buland, T. (2012). *Vurdering i Skolen. Intensjoner Og Forståelser* (Report from the project Forskning På Individuell Vurdering i Skolen). Trondheim: SINTEF
- Seelen, J. V. (2012). *Læring, praksis og kvalitet i idrætstimerne: et multiple-case studie* (Doctoral Dissertation, University of Southern Denmark, Denmark). Retrieved from <https://www.ucviden.dk/da/publications/praksis-1%C3%A6ring-og-kvalitet-i-idr%C3%A6stimerne-et-multiple-case-studi>
- Thorsnæs, E. M. S. (2019). *På hvilken måte opplever elever i videregående skole formativ vurdering i kroppsøving som læringsfremmende?* (Master's Thesis, Western Norway University of Applied Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/2605793>
- Turner, L., Johnson, T. G., Calvert, H. G., & Chaloupka, F. J. (2017). "Stretched Too Thin? The Relationship between Insufficient Resource Allocation and Physical Education Instructional Time and Assessment Practices." *Teaching and Teacher Education* 68. 210–219. doi: [10.1016/j.tate.2017.09.007](https://doi.org/10.1016/j.tate.2017.09.007)
- Utdanningsdirektoratet. (2010). *Grunnlagsdokument: Satsingen Vurdering for læring 2010 – 2014* (Report from the Norwegian Directorate for Education and Training). Retrieved

from

<https://www.udir.no/contentassets/42ddd51789024ccb8f0cea08dad9cb6e/grunnlagsdokument-satsingen-vfl.pdf>

Utdanningsdirektoratet. (2015). *Læreplan i kroppsøving (KRO1-04)* (Document from the Norwegian Directorate for Education and Training). Retrieved from <https://www.udir.no/kl06/kro1-04>

Utdanningsdirektoratet. (2018). *Erfaringer fra nasjonal satsing på vurdering for læring (2010-2018)* (Report from the Norwegian Directorate for Education and Training). Retrieved from <https://www.udir.no/tall-og-forskning/finnforskning/rapporter/erfaringer-fra-nasjonal-satsing-pa-vurdering-for-laring-2010-2018/>

Wessel, M. W. (2014). *'Var det sånn at alle skulle bli med?': en studie om prosjektet 'inkluderende læringsaktiviteter' og kroppsøvingfagets betydning for et utvalg elever* (Master's Thesis, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/221742>

Wiken, A. R. (2011). *Hva lærer vi i kroppsøving?: Hva oppfatter elevene selv at de lærer, og er dette i tråd med Kunnskapsløftet 2006?* (Master's Thesis, Norwegian School of Sport Sciences, Norway). Retrieved from <http://hdl.handle.net/11250/171703>

William, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Erlbaum.

Attachments

Attachment 1: Information sheet and consent form to participants

Attachment 2: NSD approval

Attachment 3: Interview guide (Norwegian)

Vil du delta i forskningsprosjektet

”Exploring the barriers of learning-centred assessment in Norwegian physical education classes. The teachers’ perspective”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å utforske hvilke vurderingsmetoder som brukes av kroppsøvingslærere, i tillegg til å utforske om det finnes barrierer til læringsorientert vurdering i kroppsøving. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Formålet med dette prosjektet er å utforske temaet vurdering i kroppsøving, utforske hvilke metoder som benyttes i norsk skole og lytte til kroppsøvingslæreres syn på vurdering i kroppsøving og eventuelle barrierer for å implementere læringsorientert vurdering i kroppsøvingsfaget. Prosjektet vil utføre om lag 12-20 intervjuer med kroppsøvingslærere med forskjellig bakgrunn og fra forskjellige skoler i Norge ved bruk av kvalitative semi-strukturerte forskningsintervjuer. Oppgaven er en masteroppgave ved Universitetet i Oslo, innen retning Comparative and International Education.

Oppgaven til ta stilling til følgende problemstillinger:

- *Hvilke vurderingsmetoder benyttes av kroppsøvingslærere og i hvilken grad klarer disse vurderingsmetodene å fange læringsmålene?*
- *Er det barrierer for å implementere læringsorienterte vurderingspraksiser i skolesystemet, og hva kreves for at lærere skal kunne benytte seg av læringsvennlige og bærekraftige vurderingsmetoder i kroppsøving?*

Hvem er ansvarlig for forskningsprosjektet?

Universitetet i Oslo (det Utdanningsvitenskapelige Fakultet) er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Prosjektet retter seg inn mot opplevelser forskjellige lærere har med vurdering i kroppsøving og ønsker å belyse perspektivene til lærere fra forskjellige skoler, med forskjellig bakgrunn og med forskjellige erfaringer. Forespørsler om deltakelse vil bli gitt til kroppsøvingslærere frem til et tilstrekkelig antall lærere er nådd, samtidig som utvalget representerer lærere med forskjellige karakteristikk som kjønn, alder, erfaring som lærer og erfaring som deltaker i idrett/fysisk aktivitet.

Prosjektet vil først og fremst rette seg inn mot forskjellige skolers administrasjon og rektorer med forespørsel om aktuelle kandidater til prosjektet og om eventuelt deres kontaktopplysninger. Prosjektet vil også ta imot anbefalinger fra deltakere om andre potensielle kandidater (såkalt snowball sampling).

Hva innebærer det for deg å delta?

Datainnsamlingen utføres som kvalitative semi-strukturerte forskningintervjuer med deg som deltaker, knyttet opp mot problemstillingene prosjektet ønsker å belyse. Intervjuene vil ta ca. 30-60 minutter å gjennomføre. Data innsamles ved hjelp av opptak (mobilapp Dictaphone og lydopptaker). Svarene dine fra intervjuet vil bli lagret i Universitetet i Oslos nettjeneste 'Nettskjema' som lagrer forskningsdata på en trygg måte.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykke tilbake uten å oppgi noen grunn. Alle opplysninger om deg vil da bli anonymisert. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- Det er kun student og veileder som vil ha tilgang til dine svar i løpet av prosjektet. Svarene og all personlig data knyttet til forskningsintervjuene vil slettes umiddelbart etter prosjektets slutt.
- Dine svar vil tas opp med mobilappen 'Dictaphone' som laster dataene direkte opp til UiOs tjeneste 'Nettskjema' for lagring av forskningsdata. Det vil også bli benyttet lydopptaker som reserveløsning, hvor dataene vil bli lastet opp til Nettskjema og slettet fra lydopptakeren så raskt som mulig etter intervjuene har funnet sted. Dataene knyttet til intervjuet vil ikke lagres på privat datamaskin eller mobil, kun i 'Nettskjema'.
- Det vil bli benyttet transkripsjonsverktøy (f4transkript) i forbindelse med transkriberingen av dine svar.

All data vil bli anonymisert (navn, navn på skole, navn på sted, og så videre) i prosjektets publikasjon. Spesifikk informasjon som kan knyttes til deltaker (stedsnavn, navn på tidligere idrettslag, tidligere skoler, og så videre) vil ikke bli inkludert i publikasjonen. Det vil ikke være mulig å identifisere deg som deltaker i publikasjonen. Deltakere vil bli beskrevet i oppgaven ut i fra kriterier som for eksempel:

-mann/kvinne

-alder i 40-årene

-tidligere idrettsaktiv

-10-15 års erfaring som lærer

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal etter planen avsluttes 1. August 2020. Etter dette vil alle opptak og data som kan knyttes til deg som deltaker bli slettet.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,

- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Universitetet i Oslo har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Universitetet i Oslo ved Rogers Kaliisa (rogers.kaliisa@iped.uio.no)
- Vårt personvernombud: Maren Magnus Voll (personvernombud@uio.no)
- NSD – Norsk senter for forskningsdata AS, på epost (personverntjenester@nsd.no) eller telefon: 55 58 21 17.

Med vennlig hilsen

Rogers Kaliisa
Prosjektansvarlig
(Forsker/veileder)

Thomas Sandholt Lund
Student

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet ”Exploring the barriers of learning-centred assessment in Norwegian physical education classes. The teachers’ perspective”, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. August 2020.

(Signert av prosjektdeltaker, dato)

NSD NORSK SENTER FOR FORSKNINGSDATA

NSD sin vurdering

Prosjekttittel

Exploring the barriers of learning-centred assessment in Norwegian physical education classes.
The teachers' perspective

Referansenummer

468040

Registrert

10.09.2019 av Thomas Sandholt Lund - thomslu@student.uv.uio.no

Behandlingsansvarlig institusjon

Universitetet i Oslo / Det utdanningsvitenskapelige fakultet / Institutt for lærerutdanning og skoleforskning

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Rogers Kaliisa, rogers.kaliisa@iped.uio.no, tlf: 22845318

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Thomas Sandholt Lund, thomas.s.lund@gmail.com, tlf: 96041316

Prosjektperiode

23.09.2019 - 31.07.2020

Status

11.09.2019 - Vurdert

Vurdering (1)

11.09.2019 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med

personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg 11.09.2019. Behandlingen kan starte.

MELD VESENTLIGE ENDRINGER

Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

<https://meldeskjema.nsd.no/vurdering/5d776758-ab6e-4944-9045-1b7ca049383e>
9/16/2019Meldeskjema for behandling av personopplysninger

1/2

https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html

Du må vente på svar fra NSD før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 31.07.2020.

LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER

NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om ogsamtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte ogberettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante ognødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for åoppfylle formålet

DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER

NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Dersom du benytter en databehandler i prosjektet må behandlingen oppfylle kravene til bruk av databehandler, jf. art 28 og 29.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET

NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Tlf. Personverntjenester: 55 58 21 17 (tast 1)

Exploring the barriers of learning-centred assessment in Norwegian physical education classes. The teachers' perspective

1 Personlige erfaringer (for sammenligning)

- kjønn
- alder
- hvordan mange år har du vært lærer? Og som kroppsøvlingslærer?
- har du gjennomført lærerutdanning? Har du annen relevant skole/utdanning?
- hva lærte denne utdanningen/bakgrunnen deg om vurdering?
- hva er din idrettsbakgrunn? Har du noen bakgrunn fra konkurranseidrett?

2 Hva slags vurderingspraksis bruker du som lærer?

- vet elevene hva som er forventet av dem?
- fortell om hvordan du vurderer studenter i kroppsøving? Eksempler?
- har du eksempler fra f.eks ballspill, friidrett, svømming, friluftsliv eller egentrening?
- hvordan mange aktiviteter vurderes elevene i?
- bruker du tester for vurdering eller 'integret' vurdering?
- kjenner du til tilfeller hvor fysiske tester brukes til å sette karakter?
- brukes ulike vurderingsmetoder? Egenvurdering, medstudentvurdering, prosjektarbeid, e.l.?
- er det en begrunnelse bak den typen vurdering du bruker?
- hvilke kriterier legges til grunn for vurderinga?
- innsats, holdning, ferdigheter? hvordan vektlegges disse?
- involveres elevene i vurderingsprosessen? Hvordan?
- i hvilken grad foregår læring i kroppsøving?
- Motoriske ferdigheter? Bevegelsesglede? Kunnskap om idretter? Fair play?
- føler du vurderingen fanger opp læring hos eleven? Eller mest status quo-evaluering?

- gis elevene feedback på hvordan de kan forbedre seg og karakteren sin? (formativ vurdering)
- gis elevene en karakter hvert semester (8-10)?
- føler du at vurderingen fanger kompetansemålene (ferdigheter i lagidretter, svømmeteknikker, orientering med kart og kompass)?
- har du noe inntrykk av hva elevene føler om vurderingen? → Blir de motivert av det?
- er det noen elever som sliter mer enn andre fra et vurderingsperspektiv?

3 Hva synes du er god vurderingspraksis i kroppsøving?

- hvordan synes du god vurderingspraksis ser ut i kroppsøving?
- hvor mye bør innsats/deltakelse ha å si etter din mening? Hvor mye har ferdigheter å si?
- hvordan opplevde du kroppsøving/gymnastikk i skolen? Hvordan opplevde du vurderingen?
- er det elementer av vurderingspraksisen som ble brukt når du var elev som du også bruker som lærer i dag?
- hva ser du på som formålet med kroppsøving?
- hva tenker du vurdering i kroppsøving skal måle? Hva verdsetter du i kroppsøvingfaget?
- hva ser du på som læring i kroppsøving? Hva tenker du elever skal få ut av det? Eksempler?
- har du noen tanker rundt 'vurdering for læring' i kroppsøving?
- hvordan kan vurdering i timene dine maksimere læring i faget?
- hva synes du personlig om bruken av karakterer i kroppsøving?
- føler du vurdering i kroppsøving skal være lik eller ulik andre fag? F.eks ferdigheter vs innsats
- har du noen tanker rundt Fagfornyelsen fra et vurderingsperspektiv?

4 Er det begrensninger hva slags vurderingspraksiser du bruker i kroppsøving?

- legger du noen plan for hvordan vurderinga skal foregå?
- designer du vurderingen slik du selv synes er best? Hvilke hensyn må du ta når du legger rammene for vurdering?
- har tidsressurser noen betydning for utformingen av vurdering i timene dine?
- har fasiliteter og geografi noen betydning?

- føler du at du har muligheten til å fremme læring i kroppsøving? Begrenser vurderinga deg?
- føler du at du har handlingsrom når det gjelder hvordan vurderingspraksis du skal bruke?
- hva tenker du om kompetansemålene? Er det lett å vurdere utfra disse?
- er det noen former for hjelp eller støtte du mangler når det kommer til å skape god vurderingspraksis?
- er dokumentasjon en viktig del av vurderingsarbeidet (ift klagesaker)?
- hva føler du skolen og utdanningssystemet forventer av deg når det gjelder vurdering?
- hva føler du elever og foreldre forventer av deg når det gjelder vurdering?
- er det noe du føler du kunne gjort bedre når det gjelder vurdering?
- hva føler du er den største utfordringa til kroppsøvingfaget? I forhold til vurdering?
- noe annet du vil legge til om temaet?