Avhandlingsserie för
Gymnastik- och idrottshögskolan

Nr 08

# Grading in physical education

GIH THE SWEDISH SCHOOL OF SPORT AND HEALTH SCIENCES

# Grading in physical education

Lena Svennberg

© Lena Svennberg
Gymnastik- och idrottshögskolan 2017
ISBN 978-91-980862-9-4

Tryckeri: Universitetsservice US-AB, Stockholm 2017
Distributör: Gymnastik- och idrottshögskolan

To my family,

Mother, Lasse, David, Jacob and Elsa.

# Abstract

In the thesis the aim is to investigate different aspects of what teachers value when grading in Swedish physical education (PE) and to analyses how sociological background factors impact students' grades. Grades in PE have included aspects other than those prescribed in the grading criteria, for instance motivation and effort. Teachers sometimes find their value-setting difficult to articulate and refer to a "gut feeling". In order to explore both explicit and implicit forms of value-setting, the Repertory Grid interview technique is employed.

The thesis includes four sub-studies, three interview studies with Swedish PE teachers and a fourth study based on registry data from the Swedish National Agency for Education. The data of all students leaving nine-year compulsory school in 2014 (n=95317) is analysed to explore how sociological background factors, such as migration background, parents' education, school provider and gender, affect PE grades.

The results reveal aspects of grading that are not detectable in the official description of the grading assignment and highlight problems that teachers need to address when grading. Four themes are discerned in the teachers' grading practices: motivation, knowledge, confidence and social skills. The implementation of a new national curriculum with specified knowledge requirements seems to improve the alignment with the national criteria, but there is still a gap between policy and practice. The knowledge requirements for movement are often interpreted as performances in competitive sports, even if the teachers try to find other interpretations. The odds ratio for getting a higher grade in PE is greater for the variables migration background and parents' education than for the other investigated variables. The concepts formulated by Bernstein are applied to explore the relations between teachers' grading practices and cultural and political influences and to discuss how the tensions between different interests could affect teachers' grading.

The conclusion is that the gap between policy and practice confirmed in this study is related to tensions between the interests and purposes of different agents, all of whom strive to influence steering documents and practice. Cultural and political influences need to be considered and facilitate discussions about how to understand which knowledge is valued in PE and who has better possibilities to assimilate it.

# Articles

I Svennberg, L., Meckbach, J., & Redelius, K. (2014). Exploring PE Teachers' 'Gut Feelings': An Attempt to Verbalise and Discuss Teachers' Internalised Grading Criteria. *European Physical Education Review*, *20*(2), 199-214. doi: 10.1177/1356336X13517437

With permission from Sage

II Svennberg, L., Meckbach, J., & Redelius, K. (published online 2016). Swedish PE teachers struggle with assessment in a criterion-referenced grading system. *Sport, Education and Society.* doi: 10.1080/13573322.2016.1200025

With permission from Taylor & Francis

III Svennberg, L. (published online 2016). Swedish PE teachers' understandings of legitimate movement in a criterion-referenced grading system. *Physical Education and Sport Pedagogy.* doi: 10.1080/17408989.2016.1176132

With permission from Taylor & Francis

IV Svennberg, L. & Högberg, H. Who gains? Sociological parameters for obtaining high grades in physical education. Submitted.

# Contents


1. Introduction ..... 15

Aim of the thesis ..... 18

Reading guidelines ..... 19

2. Background ..... 20

The Swedish school system ..... 20

PE teachers' grading practices ..... 23

The value of motivation and effort when grading ..... 24

Valued movement in PE ..... 26

Equity in education and PE ..... 27

Understanding teachers' grading practices ..... 29

3. Theoretical framework ..... 30

The interrelated message systems ..... 31

The instructional- and the regulative discourse ..... 31

Visible and invisible pedagogy ..... 32

The pedagogic device – part one ..... 33

The competence model and the performance model ..... 34

The pedagogic device – part two ..... 35

4. Methodology ..... 37

The Repertory Grid technique ..... 37

How the interviews are performed ..... 39

Selection ..... 41

Data collection ..... 42

Analysis ..... 43

Methodological considerations ..... 45

Ethical considerations ..... 47

5. Summary of articles ..... 49

Article I: Exploring PE teachers' 'gut feelings': An attempt to verbalise and discuss teachers' internalised grading criteria ..... 49

Article II: Swedish PE teachers struggle with assessment in a criterion-referenced grading system 50

Article III: Swedish PE Teachers' Understandings of Legitimate Movement in a Criterion-referenced Grading System ..... 51

Article IV: Who gains? Sociological parameters for obtaining high grades in physical education .. 53

Summary overview ..... 55

6. Discussion ..... 56

Critical considerations ..... 56

Equity in PE ..... 57

Agents in the transformation of policy into practice ..... 60

The relations between the micro- and macro levels ..... 62

Contextualising three dilemmas in the teachers' grading practices ..... 64

The advantages and disadvantages of the models ..... 66

Further questions ..... 68

7. Conclusions and implications ..... 70

8. Summary in Swedish ..... 72

Bakgrund till studien ..... 72

Syfte och frågeställningar ..... 73

Teoretiskt ramverk ..... 74

Metoder ..... 77

Sammanfattning av artiklarna ..... 78

Artikel I: Exploring PE teachers' 'gut feeling': An attempt to verbalise and discuss teachers' internalised grading criteria ..... 78

Artikel II: Swedish PE teachers struggle with assessment in a criterion-referenced grading system ..... 79

Artikel III: Swedish PE Teachers' Understandings of Legitimate Movement in a Criterion-referenced Grading System ..... 80

Artikel IV: Who gains? Sociological parameters for obtaining high grades in physical education ..... 81

Diskussion ..... 82

Lika villkor i idrott och hälsa ..... 82
Aktörer i transformeringen från läroplan till praktik ..... 84
Prestationsmodellen och kompetensmodellen ..... 85
Kontextualisering av tre dilemman från studien ..... 86
Slutsats och implikationer ..... 88
Acknowledgements ..... 90
References ..... 92

# 1. Introduction

*"I tear my hair out to give students extra practice to reach the knowledge requirements for the grade C in dance. These are students, who have an A in everything else, and I would like to give them a B instead of a D"*

*"I wish the National Agency for Education could give examples of how to assess risks with physical activities, so we can do it in an equivalent way. I miss that, you always have to reinvent the wheel yourself…"*

*"I have an obese student, who is well able to adapt his/her movements, but cannot perform all the movements due to obesity. Can this student still get the highest grade in the knowledge requirement for movement?"*

*(Quotes from a Facebook group for Swedish teachers in Physical Education and Health [April 2016, author's translation])*

This thesis is about grading in the Swedish school subject physical education and health (PE)[1]. Grading students in school subjects is difficult because many dimensions need to be accounted for. The above quotations are examples of typical situations that PE teachers face in their grading practices. The PE teachers quoted above use social media to share their problems with teaching and grading. For a few weeks in April 2016, as a typical example, the grading issues they discuss as most problematic are how to interpret the national grading criteria and how to find equity and equivalence in their grading practices. Grading students in PE can be more complicated than grading in other school subjects, in that the physical body of the learning student is involved in processes of socialisation and identity formation (Lundvall, 2014). Here, success or failure in students' physical performances is made visible in a special way and at a sensitive time in

[1] The Swedish name for the subject, idrott och hälsa, is translated as physical education and health (PEH). In this thesis, and in the published articles that are included, the name is instead translated as physical education (PE), which is recognised internationally.

their lives. In order to address the questions the teachers ask, it is important to investigate what they value in their grading work.

A problem when investigating teachers' grading practices is that they sometimes find it difficult to articulate what they value and occasionally rely on their gut feelings (Annerstedt & Larsson, 2010; Hay & MacDonald, 2008; Swedish National Agency for Education [SNAE], 2000). These gut feelings can be regarded as tacit knowledge that is part of PE teachers' professionalism and something that is passed on within the profession. For instance, without necessarily being able to express movement qualities in words, the tacit dimension can enable teachers to make professional assessments of them. However, the lack of transparency in the tacit dimension can be problematic in a goal and criterion referenced grading system, such as that used in Sweden.

Transparency is necessary if students are to know what they are supposed to learn. In Sweden, students are entitled to know what is expected at the different grading levels and teachers are held accountable for the grades they set. Clearly communicated learning objects have also been shown to improve learning (Hattie, 2012). Another potentially problematic aspect is that transparency is needed in order to evaluate the equivalence of grades. In this context, the tacit dimension could be regarded as a threat to validity and reliability. To quote Annerstedt and Larsson (2010):

> …the principles of fair and equitable grading in Swedish PE should be strongly questioned due to lack of transparency, validity and reliability. Grading in Swedish PE seems to be arbitrary and the grading system is not accessible or transparent to students (p. 97).

A point of departure in the thesis is that grading is regarded as an important message system that defines and communicates to students which knowledge they need to learn (Bernstein, 2003a). In Sweden, specific national grading criteria, also called knowledge requirements, are prescribed as the only grounds for PE grades. However, similar to other school subjects, a discrepancy has been identified in Swedish PE between the national grading criteria and the actual grading that takes place in schools (Annerstedt & Larsson, 2010; Eriksson, Gustavsson, Quennerstedt, Rudsberg, Öhman & Öijen, 2005; Redelius, Fagrell & Larsson, 2009; Tholin, 2006). There is therefore a risk that when the grading practice does not align with the official grading criteria, it sends conflicting messages to students about what is important to learn. As the process of assessment and grading constitutes a strong message system (Bernstein, 2003a; Hay & Penney, 2013), it has the potential to affect students' learning. Therefore, it is important to get as full a picture as possible of what teachers value when setting grades in PE.

This overall picture could help to promote teachers' professional conversations about the evaluation and development of their own grading practices. Exploring what teachers value in different contexts could also contribute to discussions about equivalent grading,

and make it possible to make use of their professional experiences of applying different grading strategies in different contexts. Consequently, a central question in the thesis is what PE teachers 'assign value' to when grading students in the school subject of PE.

In addition to the goal of equivalent grading, the Swedish school system is based on the general principle of equity. The Swedish National Agency for Education describes different aspects of equity and states that "The concept of equity can be divided into three fundamental aspects: equal access to education, equal quality of education and the compensatory nature of education" (SNAE, 2014, p. 7). In this thesis, the focus is on the compensatory nature of education. The steering documents make it clear that equity in education does not mean that all education is alike, but rather that "education must be adapted to students' different circumstances at home and other circumstances such as having Swedish as a second language, gender, possible disabilities etc." (SNAE, 2014, p. 9). The national curriculum imparts the message that while all learning is to be promoted, an individual's background and specific needs should be taken into account. Thus, in the educational context, equity means providing all students with whatever they need to learn successfully. When the learning process is based on an individualised pedagogy, which allows for multiple ways of acquiring and expressing knowledge, learning outcomes are more difficult to compare (Gipps, 1994).

Since the middle of the 1990s, when both the goal referenced grading system and free school choice were implemented in Sweden, equity, assessed on the basis of results from different categories of pupils, has deteriorated (SNAE, 2006, 2014). The PE grades of different groups of students leaving nine-year compulsory school are published annually by SNAE. However, even though we know that PE grades may vary with different sociological variables, less is known about how different contextual factors interact. Therefore, another question posed in the thesis is how different sociological variables impact PE grades both separately and together.

It is important to explore which criteria teachers use in their grading practices as a first step to understanding what they value in their grading decisions and which factors contribute to this value-setting. One way of identifying patterns, that could facilitate a discussion about influencing factors is to analyse the messages transmitted in PE teachers' grading practices in relation to those that can be discerned in the national grading criteria and in the overall curriculum. In order to achieve a better understanding of teachers' grading practices, it is also important to contextualise the results of this thesis in cultural, sociological and institutional terms. Sicilia-Camacho and Brown (2008) recognise a need to "connect the practical, procedural and technical aspects of everyday pedagogical practice with broader historical and political discourses; or put in more sociological terms, the connections between the micro to the macro" (p. 103). The intention of the thesis is thus to contribute to the verbalisation and problematisation of teach-

ers' value-setting when grading in relation to wider contexts and to identify the possible equity consequences of their value-setting.

## Aim of the thesis

The general aim of the thesis is to explore the grading practices of Swedish PE teachers in the final year of nine-year compulsory school. A special focus is on what PE teachers value when setting grades in relation to the specified grading criteria and the national goals for students' overall education in PE. Further aims are to analyse and problematise the distribution of grades among different groups of students and to explore whether and how sociological background factors impact PE grades. In this exploration of the grading practices in Swedish PE, four research areas are investigated. Each area is animated by four separate studies, resulting in four research articles. These research areas are:

- Which values PE teachers consider important when grading and how these values affect the grades they set (reported in article I: *Exploring PE Teachers' "Gut Feelings": An Attempt to Verbalise and Discuss Teachers' Internalised Grading Criteria*).
- The impact of the implementation of national knowledge requirements on what PE teachers consider important when assigning grades, and which non-knowledge related aspects (if any) teachers continue to address after the implementation (reported in article II: *Swedish PE teachers struggle with assessment in a criterion-referenced grading system*).
- Which movements PE teachers consider legitimate to grade in a criterion-referenced grading system, and which factors influence their grading (reported in article III: *Swedish PE teachers' understandings of legitimate movement in a criterion-referenced grading system*).
- Whether and how sociological background factors impact PE grades (reported in article IV: *Who gains? Sociological parameters for obtaining high grades in physical education.*).

To explore the fourth research area, national data from the Swedish National Agency for Education has been utilised. The data is analysed in order to explore how four different background factors, namely gender, parents' educational background, migration background and attendance at municipal or independent schools, impact PE grades separately and together.

An analytical aim is to analyse and problematise teachers' grading practices at the micro level, in the classroom, in relation to the macro level national curriculum and influences on the Swedish school system. A curriculum theoretical perspective and Bernstein's concept of pedagogic device, as outlined in Chapter 4, are used to analyse

what teachers value in their own pedagogical and didactic practices in the light of steering documents and agents striving for influence over education policy and practice.

## Reading guidelines

The thesis is disposed as follows: Chapter 2 begins by describing the Swedish school system. Thereafter, the research field is outlined. The reader is presented with an analysis of previous research on teachers' grading practices, an overview of links between school grades, differences in students' backgrounds and an understanding of the gaps between policy and practice. In Chapter 3, theoretical concepts, garnered from curriculum theory and outlined by the British sociologist Bernstein (1996, 2003a, 2003b), are explored to determine how they might apply to the Swedish school system. The methodological approaches used in each of the four studies are described and argued for in Chapter 4. Summaries of the four studies included in the thesis are presented in Chapter 5. In Chapter 6 follows an elaborated discussion of how Bernstein's conceptual framework, presented in Chapter 4, can be used as a tool to interpret and understand the results of the four studies. The results of what teachers value when grading are analysed in relation to the more overarching discourses proposed by Bernstein. In this chapter, critical considerations and suggestions for further research are also put forward. Finally, in Chapter 7, the conclusions and implications for this thesis are summarised.

# 2. Background

In this chapter, research on grading from teachers' perspectives and what PE teachers value when assigning grades is outlined first. Thereafter, studies and reports about equity are presented. National and international research published in peer-reviewed scientific journals is also reported, together with doctoral theses and national and international reports and evaluations. The focus is on the Swedish context, although international research is also included. Comparisons between different countries can be problematic, however, because grading has different functions in different nations and works differently in different contexts. With this in mind, the results from different countries are sometimes presented together in order to illustrate the similarities between them. As this thesis focuses on the final year of Swedish nine-year compulsory school, research on grading in higher education has been excluded.

The concepts of grading and assessment are closely related and in some studies are used synonymously. However, there is a reason for making a distinction between them, because aspects that are assessed are not always given the same value when grading (e.g. Eriksson et al., 2005). Hay and Penney define assessment as "any action of information collection within education settings that is initiated for the purpose of making same interpretive judgements about students" (2013, p. 6). While assessment involves the collection and interpretation of information, "grading represents a summary of that interpretation, communicated via a symbol" (Hay & Penney, 2013, p. 8). In this thesis, the assessments that teachers attach value to when summarising them in a grade are explored. Before starting the presentation of previous research, the Swedish school system is described in brief, including the grading system, to contextualise the research areas.

## The Swedish school system

The Swedish school system has changed from being one of the most regulated at the beginning of the 1990s to one of the most deregulated in the western world (Duan, 2006; von Greiff, 2009). A free school choice and competition for students between schools have led to an education market that is supposed to be more effective and of

higher quality (Lindblad, Lundahl, Lindgren & Zackari, 2002; Samuelsson & Lindblad, 2015). A new situation emerges when local authority schools compete for applicants with private but publically founded independent schools. A "New Public Management" decentralises responsibility and organisation to municipalities, while the state claims control by stipulating goals to achieve. The decentralisation of government over schools is combined with an increased centralised control from the state over school results, for instance by means of students' grades and national tests (Sundberg & Wahlström, 2012).

A free market provides choices for individuals, who are in turn held responsible for the consequences of their choices. Responsibility for the economy and human well-being are transferred from the state to a market in which people are considered as entrepreneurs of their own lives and where personal responsibility is emphasised (Davies & Bansel, 2007). The free market is not just a national issue. Globalisation opens up possibilities for new markets for school providers. In a school context, globalisation is visualised in international testing, where the outcomes of different countries' school systems are compared.

Before 1994, the ideology of the Swedish school system was one of public welfare (Martinsson & Reimers, 2014). Publicly financed schools were intended to give everyone, regardless of background, an equal (interpreted as the same) education and to serve as meeting places for different social and cultural groups (Martinsson & Reimers, 2014). Schools tried to compensate for class inequalities and educated students for different roles in society. Grades were norm referenced and compared students who had received what was supposed to be the same education. The state regulated how schools should work in a national curriculum.

A reformed curriculum for the nine-year compulsory school, called Lpo 94, was implemented in 1994. It introduced the present goal referenced grading system. Grades have several functions in the Swedish school system. They are used as a selection instrument for the next education level and are intended to inform students and parents at which level the grading criteria have been accomplished. The motivational effect of grades is sometimes emphasised, even though this effect has been questioned by scholars (Lundahl, Hultén, Klapp, Mickwitz, 2015). The introduced goal referenced grades are also used to evaluate schools and are regarded as indicators in the follow-up and reconciliation of the stipulated goals (Swedish National Audit Office, 2011). School grades, on a group level, are published and made available to students and parents when deciding which school to choose. In this sense, they can be regarded as marketing tools.

A problem with Lpo94 was that teachers experienced the grading criteria as too vague and too difficult to interpret (Selghed, 2004; Tholin, 2006). As a result, a new

national curriculum was introduced in Sweden in 2011, called Lgr11 (SNAE, 2011). The first two parts of the overall curriculum, *fundamental values and tasks of the school* and *overall goals and guidelines* for education are basically the same as those of the former curriculum. School education in Sweden has two goals: to "promote learning by stimulating the individual to acquire and develop knowledge and values" (SNAE, 2011, p. 11) and "to promote values and norms in order to foster democracy" (SNAE, 2011). The major changes in the latter reform are visible in the third part of the curriculum, namely the syllabuses for the different subjects. The syllabus in Lgr 11 has the same structure for each subject and consists of three parts: aim, core content and knowledge requirements.

In the aims stipulated for PE, teachers should encourage students to develop knowledge in different areas. Non-knowledge aspects are also included, such as developing an interest in being physically active, spending time outdoors, a healthy lifestyle, interpersonal skills, respect for others and a belief in their own capacities. The second part of the PE syllabus, core content, is organised in three knowledge areas: *movement, health and lifestyle* and *outdoor life and activities*.

Finally the third part of the PE syllabus, knowledge requirements, focuses on what students are expected to know or are able to perform. The new specified knowledge requirements determine what to grade and have a progression between the different grade levels. Sundberg and Wahlström (2012) write that these requirements claim to be standardised, measurable and comparable and codify a specific conception of knowledge. Regulations pertaining to who is allowed to grade in terms of teacher registration also contribute to the construction of the professional teacher as someone who is able to follow the regulations relating to knowledge requirements (Mickwitz, 2015). The grading scale in Lgr11 has increased from three to six levels: A-F. The highest grade is A and the lowest F, which designates failure. Knowledge requirements are prescribed for grades A, C and E. The in-between grades are used when some, but not all, of the requirements for the higher grade are fulfilled. SNAE provides support for interpreting the requirements on its website.

In a comparison of Lpo94 and Lgr11, Wahlström (2014) found that the focus has shifted from "doing" in Lpo94 to "knowing" in Lgr11, and that in Lgr11 the knowledge that is valued and to be assessed has more focus on instrumental values. Sivesind (2013) agrees that the assessment of students' results is in focus, but comments that some parts of curriculum also focus on aim, output and content, thereby merging "professional semantics on the purposes, and content of schooling with a language formed by evidence-based policy on outcomes" (Sivesind, 2013, p. 52). Many western countries' school systems have developed in a similar way and now include marketing, standardised grading and an increased focus on measuring (Apple, 2005). This is the context in

which the teachers in this thesis work when grading. In the following I will present research on teachers' grading.

## PE teachers' grading practices

A gap between PE teachers' grading practices and official grading policy is visible in several studies (Ekberg, 2009; Redelius et al., 2009: Tholin, 2006). In the former curriculum, Lpo94, PE teachers' interpretations of the criteria were found to differ (Tholin, 2006). For instance, in a national evaluation of PE some ten years after the implementation of Lpo94, some national criteria, such as those for dance, orienteering, swimming and outdoor activities, had less impact on the PE grade, in the sense that students with the highest grade reported that they had no or little knowledge in those areas (Eriksson et al., 2005). Another example is that interpretations of the national criteria do not always reflect their intentions. For example, the PE syllabus includes functional physical exercise and conceptual development, as well as an understanding of health and lifestyles. However, in practice, the lesson content is often dominated by competitive sport and physical exercise (Ekberg, 2009). There is also a close relation between content in the lessons and what is valued in the grades. Redelius et al. (2009) found that good performances in competitive sports were important, especially for the highest grades in PE.

In their study, Annerstedt and Larsson (2010) found that Swedish PE teachers have developed a grading practice that is based on their own experiences and that this is not always in line with the national grading criteria. Research on Swedish PE shows that there is a lack of clarity of the educational purpose and that teachers find it difficult to express what important knowledge in PE is (Larsson, Fagrell & Redelius, 2005; Larsson & Karlefors, 2015; Redelius, Quennerstedt & Öhman, 2015). This vagueness is also observed in other countries and has initiated a discussion about the possible future of the subject (Kirk, 2006, 2010; Penney & Chandler, 2000; Tinning, 2000).

Due to these blurred objectives, it can be difficult to express which criteria are used when grading. PE teachers occasionally refer to internalised grading or gut feelings (Annerstedt & Larsson 2010; Hay & MacDonald, 2008; SNAE, 2000). Gut feeling is when a PE teacher intuitively knows that a student should be assigned a certain grade, but finds it difficult to express why. The teachers in Hay and MacDonald's study (2008) do not refer to the official grading criteria in their judgements, but instead suggest that the criteria are internalised. The use of internalised criteria raises several questions, such as what teachers are actually grading and how the differences between criteria and practice could be understood in the Swedish PE context. Although several studies have

examined what teachers value when assigning grades, there are still gaps in the knowledge. One reason for this may be the difficulty of investigating the internalised criteria that the teachers themselves find difficult to articulate. Therefore, the aim in the first article is to explore which explicit and implicit personal criteria teachers participating in the study attach value to when grading. In order to develop this area, an interview technique known as the Repertory Grid (RG) technique, is used. The technique has proven useful for verbalising perceptions of something that is familiar but is difficult to express.

## The value of motivation and effort when grading

The Swedish knowledge requirements for PE prescribe knowledge as the only basis for grades. Before the introduction of the knowledge requirements and the stronger regulation of what to grade, Swedish PE teachers sometimes valued criteria like motivation and effort when assigning grades (Annerstedt & Larsson, 2010; Klapp Lekholm & Cliffordson, 2009; Redelius et al., 2009; Tholin, 2006). The goal-referenced grading system, which was introduced in Sweden in 1994, presupposes that the same knowledge will warrant the same grade, regardless of which school students attend or who their teachers are. However, the Swedish National Audit Office (2011) has questioned whether such equivalence exists. Validity is a central theme in research on teachers' grading practices in different subjects. In a review by Lundahl et al. (2015), Swedish research is identified as focusing on the relation between steering documents and practice, while international research is more interested in what teachers grade, i.e. students' skills and/or personal qualities. The impact of students' individual characteristics on grades has also been studied. Here, gender, ethnicity and socioeconomic status are identified as impacting factors, as well as students' behaviour and effort (Lundahl et al., 2015). Research on students' behaviour and effort is presented in this section. The other impacting factors are presented under the heading *Equity in education and PE*.

Already in the 1950s Parsons argued that both the school's mission to promote students' knowledge and socialisation affected grades:

> In the elementary grades these two primary components are not clearly differentiated from each other. Rather, the pupil is evaluated in diffusely general terms; a good student is defined in terms of a fusion of the cognitive and the moral components, in which varying weights are given to one or the other. Broadly speaking, then, we may say that the "high achievers" of the elementary school are both the "bright" pupils, who catch on more easily to their more strict intellectual tasks, and the more "responsible" pupils, who, "behave well" and on whom the teacher can "count" in her difficult problem of managing the class (Parsons 1959, in Cross & Frary 1999, p. 54f).

In the 1990s, Brookhart referred to grades as a "Hodgepodge grade of attitude, effort and achievement" (Brookhart, 1991, p. 36). The fusion of knowledge with motivation

and effort as a common ground for assigning grades is verified in studies of PE and other school subjects (Biberman-Shalev, Sabbagh, Resh & Kramarski, 2011; Brookhart, 2013; Chan, Hay & Tinning, 2011). In addition to motivation and effort, both of which are valued in many subjects, other aspects of being a "good student", such as changing clothes, fair play and leadership skills, are also valued in PE grades (cf. Redelius et al., 2009). An international review shows that assessment in PE has traditionally focused on non-knowledge aspects, such as results in personal fitness tests and students' characteristics, but that alternative forms of assessment facilitate a stronger focus on the educational aspect (López-Pastor, Kirk, Lorente-Catalán, MacPhail, & Macdonald, 2013). The impact of values and norms on grades varies in different situations. Teachers of mathematics seem to focus more on the knowledge aspect than language teachers (Biberman-Shalev, et al., 2011) and physical education teachers (Prøitz, 2013). Paying attention to students' characteristics when grading younger students (Chan et al., 2011), lower-achieving students (Korp, 2006; Zoeckler, 2007) and students from lower socioeconomic backgrounds (Klapp Lekholm & Cliffordson, 2008) also seems to be common. This suggests that different classroom situations influence the balance between knowledge and effort when grading. It also implies that as students from different socioeconomic backgrounds can be assessed differently, the equity aspect, studied in article IV, is affected.

The absence of official visible criteria for values and norms does not mean that they are not known to students. Sivenbring (2016) finds that "students are more familiar with the disciplinary regulations and expectations than with the knowledge requirements" (p. 257). When students find the knowledge requirements difficult to understand, they instead use the strategy of displaying generic knowledge about being a "good student" to receive higher grades (Sivenbring, 2016). This is also recognised in PE. For example, Redelius and Hay (2009) found that students perceive behaviour and characteristics as the most important for obtaining a high grade in PE; a perception that is not reflected in the official assessment expectations in the PE syllabus.

As described above, the tradition of grading the characteristics of a "good student" is recognised by teachers and students in many subjects and countries. However, this can cause problems in a grading system that strongly regulates what to grade. In turn, the grading of traits, which is not listed in the national criteria, threatens validity. Further, the message about what valued knowledge is, as transmitted by the grades, may affect students' perceptions of what to learn and draw their attention away from the official criteria. The question is, will the implementation of official knowledge requirements change the tendency among teachers to assess and grade motivation and effort and thereby improve their adherence to the knowledge requirements? This is addressed in

article II, where the aim is to explore how the implementation of knowledge requirements affects what the (same) teachers value before and after the implementation.

## Valued movement in PE

Movement qualities are to be assessed according to the knowledge requirements in PE and not on quantitative results based on performance in competitive sports. Even though the concept of movement quality is new, movement has historically been contextualised in different ways in Swedish PE. In the first half of the 1900s, Ling gymnastics formed the main content of PE. During the latter half of the century, competitive sports and activities that could be measured became more important (Larsson & Redelius, 2008; Sandahl, 2005). Sport federations contributed to shaping the subject by providing literature, equipment and in-house training courses for PE teachers (Lundvall & Meckbach, 2008), thereby influencing the content of PE. Furthermore, PE teachers (Lundvall & Meckbach, 2010) and students (Thedin Jakobsson, Lundvall, Redelius & Engström, 2012) have often been, and still are, active participants in different sport federations. In the past, the meaning of the subject was understood as developing skills in different sports and finding a suitable sport to continue to be active in outside school (Larsson & Redelius, 2008).

In 1994, the subject changed its name from *physical education* to *physical education and health.* As the name indicates, a health discourse now influences the PE syllabus at the expense of the prevalent sport and activity discourse. A national evaluation of PE shows that the educational aspect is difficult for teachers to apply in pedagogic practice (Eriksson et al., 2005) and that an activity and sport discourse is still visible. Activity is a prerequisite for the learning of movement and provides teachers with assessment opportunities. Active participation has sometimes been sufficient for a pass grade (Redelius et al., 2009; Tholin, 2006; Öhman, 2007). Studies also show that simply being active does not result in the highest grade in Swedish PE. For this, high performance in different sports is expected, even though good results in competitive sports are not prescribed in the national grading criteria (Redelius, Fagrell & Larsson, 2009). However, there is a close relation between the sports movement (e.g. The Swedish Sports Confederation with associated sport federations) and Swedish PE (Londos, 2010; Schenker, 2011) in that competitive sports influence what is taught in PE, and students with leisure time sporting experience often have higher grades in PE (Eriksson et al., 2005).

In the 2011 curriculum (Lgr11), movement is one of three knowledge areas in the PE syllabus. The grades are expected to reflect the movement qualities[2] and the require-

[2] One exception from the qualitative knowledge for movement is the ability to swim 200 metres.

ments for movement are expressed as students being able to: "vary and adapt their movement to some extent/relatively well/well to activities and context" and to "beat, rhythm and context" (SNAE 2011, p. 54f). The qualitative knowledge requirements are not related to competitive sports and measurable sport results. Instead, teachers are able to choose which movements they want to teach and assess.

The other two areas included in the Swedish knowledge requirements are *health and lifestyle* and *outdoor life and activities*. Even though all three knowledge areas are emphasised equally in the knowledge requirements, Seger's study (2015, p. 102) indicates that that there is a tendency to prioritise the knowledge requirements for movement when weighting the different knowledge requirements together for grading purposes. This therefore makes movement an important knowledge area to investigate. The new knowledge requirements for movement listed in Lgr11 (SNAE, 2011) allow for a wider concept of knowledge, thus enabling more students to feel able and included. The question is how this concept is understood in teachers' talk about how to assess the new knowledge requirements. Against the background of the dominating activity and sport discourses, the aim of article III is to explore PE teachers' interpretations of the concept of movement qualities in the knowledge requirements listed in Lgr11.

## Equity in education and PE

The results in Swedish schools differ with respect to gender, parents' education, immigration status and which school students attend (SNAE, 2014). The effect of social and cultural background on grades must be considered in the light of who has the power to decide and control what should be valued. In that sense, the groups that are advantaged or disadvantaged in the setting of grades tell us something about the origin of the prioritised knowledge. It should also be recognised that "children from lower socioeconomic areas are less likely to have access to the same social and cultural resources as their higher socio-economic counterparts" (Hay & Penney, 2013, p. 3).

Grades are sometimes used to evaluate equity. In a comparison of equity aspects from the late 1990s to 2011, SNAE concludes that educational equity has deteriorated in Swedish schools due to the fact that school choice and migration background have become more important over time. The importance of students' socioeconomic backgrounds remains high (SNAE, 2014). The growing differences in school results have been explained by an increased segregation in society, leading to a concentration of students with the same background in the same schools (the Swedish Schools Inspectorate, 2010:14). The results of students born abroad have deteriorated over time. The lowest results are now found amongst those who arrived in Sweden after the usual

school starting age (SNAE, 2014). At group level, it has been found that girls, students with at least one parent with a post-upper secondary school education qualifications and students with a Swedish background have higher grades.

PE differs from other subjects in that it is the only subject in Sweden and in many other countries in which boys have higher grades than girls.[3] This has been described as being due to the embeddedness of masculine values, such as strength, speed and controlled aggression, in the sport discourse (Kirk, 2010; Redelius et al., 2009). Another possibility is that the content of PE lessons is seemingly dominated by ball games (Eriksson et al., 2005) and that as boys tend to be more active in sport federations (Elofsson, Blomdahl, Lengheden & Åkesson, 2014; Norberg, 2016) they may benefit from this experience. A national evaluation of Swedish nine-year compulsory school (Eriksson et al., 2005) showed that sporting experience outside school is a strong predicting factor for PE grades.

In a research context, gender studies are a common way of approaching equity in PE (Davis, 2003; Larsson, Quennerstedt & Öhman, 2014; Oliver & Kirk, 2016). Migration background and ethnicity are also areas that have been well studied (Hansen, 2014; Lundvall & Safizadeh, 2012). However, less is known about social class in relation to PE. Evans and Davies (2008) draw our attention to the "shadowy presence" of social class in PE research. They call for a more central place in research to promote the understanding of "what PE is for, what achievements and 'abilities' are to be valued and what we as researchers and teachers can reasonably be expected to achieve" (p. 199).

What PE teachers expect students to perform can differ. Expectations can be regarded as appropriate expectations or self-fulfilling prophecies (Hay, 2012, p. 318). Research shows, for example, that teachers do not expect girls to perform exercises requiring a lot of strength (Larsson, Redelius & Fagrell, 2009), which reproduces the norm of strong boys and weak girls and what is expected of them.

Socioeconomic, ethnic and gender inequalities in academic achievement have been widely reported and reports on equity at a general grade level are published by SNAE (2006, 2014). Much less is known about how sociological background factors affect equity in PE achievement, i.e. PE grades. In particular, few studies have been conducted with an intersectional approach in relation to PE grades. Article IV aims to identify and analyse how sociological background factors affect PE grades, separately and together.

[3] In 2016, for the first time girls had slightly higher grades than boys in PE when leaving nine-year compulsory school. However, this is not the case for all groups of girls. For instance, girls with foreign backgrounds still have lower grades than boys with foreign backgrounds (SNAE, 2016; SNAE's website).

# Understanding teachers' grading practices

One of the overarching analytical aims in the thesis is to analyse teachers' grading practices, as revealed in the articles, in relation to curriculum and agents striving to influence education at the curriculum level and in practice. Earlier research has discussed the complex web of factors influencing grading, where the focus is often on the teacher and the local context. McMillan and Nash (2000) find that teachers strive to find a balance between their own beliefs about education and learning and student or classroom characteristics and external factors. These tensions have led MacMillan and Nash (2000) to suggest that teachers might consider grading as a fluid set of principles that change when conditions change, and that this could explain the variations in teachers' grading practices. Teachers' beliefs and values have also been found to influence what is considered valued knowledge (Hay & MacDonald, 2010). Teachers' knowledge of the subject and pedagogy is also thought to influence grading (Linde, 2012; Selghed, 2004). Moreover, it has been observed that teachers try to adapt their grading practices to the classroom situation (Brookhart, 1994; McMillan, 2003).

Teachers are aware of the importance of student and teacher relations, for instance in order to motivate students. Rinne (2014) studied grade conferences between students and teachers in upper secondary school and found that there was a tension between the official task of grading and the relational dimension. The relational aspects sometimes affect teachers' grading, for example in attempts to meet students' expectations.

The research described above focuses on the teacher, the classroom situation and the local context. Teachers have sometimes been discussed as ignorant or unwilling to follow the prevailing regulations (Mickwitz, 2015). In order to understand teachers' grading practices better, there is a need to connect aspects of everyday grading practice with broader historical and political discourses (Sicilia-Camacho & Brown, 2008). Factors that have been addressed in Swedish research in order to explain teachers' grading practices are political decisions (Klapp Lekholm & Cliffordson, 2008; Korp, 2006), parents' and principals' influence (Mickwitz, 2011), school culture and school facilities (Linde, 2012). The conceptual framework developed by Bernstein (1996, 2003a, 2003b) helps to reveal the connections between teachers' grading practices and the macro level of curriculum and socio-political decisions. In this sense, the framework can provide new perspectives and broaden the understanding of teachers' grading practices.

# 3. Theoretical framework

Curriculum theory, inspired by the British sociologist Basil Bernstein, is employed in this thesis to analyse teachers' grading practices in relation to curriculum and agents striving for influence over curriculum and practice. Bernstein's (1996, 2003a, 2003b) conceptual framework has also helped me to analyse how the results affect equity. Bernstein was engaged in equity in education throughout his working life and began by studying how language impacts success in school for children from different social classes. He was also interested in the relation between education and class reproduction and introduced the concepts of classification and framing to analyse pedagogic practice and curriculum. Classification is how a school subject is demarcated in relation to other subjects. A strong classification indicates that a subject is well specified, easy to recognise, has a strong identity and that it is clear what is part of the subject and what is not (Bernstein, 2003a). Swedish PE has been recognised as a subject with weak classification (Ekberg, 2009; Kougioumtzis, 2006; Lundvall & Meckbach, 2008), in that it is open to influences from various agents and different interests may affect which knowledge is considered legitimate. This can lead to differences in the selection of legitimate knowledge (Ekberg, 2009).

Framing is the principle of how the category is regulated and the communication that takes place between transmitter and acquirer. Strong framing indicates that the transmitter (the teacher) sets the rules for communication, whereas weak framing indicates that the acquirer (the student) has a greater degree of influence over what, how, when and where knowledge is to be transmitted and received. In this thesis, the concept of classification is applied when studying what is valued by teachers, i.e. what is legitimate to grade in PE. The national knowledge requirements indicate a strong classification, in the sense that they specify the criteria to be included in the grade and exclude all other criteria. On the other hand, the content of the requirements, for instance *to be able to adapt movements,* indicates a weak classification that is open to diverse interpretations of what to assess. The grade levels are often measured qualitatively, e.g. *to some extent/relatively well/well,* which also results in a weak classification of the different grade levels.

## The interrelated message systems

In order to shed light on the relation between curriculum and practice, as discussed in the first two articles, I turn to the interrelated message systems of curriculum, pedagogy and evaluation proposed by Bernstein: "Curriculum defines what counts as valid knowledge, pedagogy defines what counts as valid transmission of knowledge, and evaluation defines what counts as valid realization of this knowledge on the part of the taught" (Bernstein, 2003a, p. 85). Evaluation, which Hay & Penney (2013) call assessment, is in focus in this thesis. The messages that are transmitted by teachers' assessments and grading are analysed in the first article. In the second article, the messages of assessment and grading are compared to those transmitted by the overall curriculum and knowledge requirements (Lgr11). Penney, Brooker, Hay & Gillespie (2009) present and explore the links between curriculum, pedagogy and assessment as a key to overall quality PE, although quality "always needs to be contextualised in cultural, social and institutional terms" (Penney et al., 2009, p. 421). When curriculum, pedagogy and assessment send the same message, students' learning experiences can be improved.

## The instructional- and the regulative discourse

In article II, Bernstein's (1996) concept of a pedagogic discourse is employed to analyse the interrelated message systems mentioned above. The pedagogic discourse consists of an instructional discourse, which creates specialised skills (knowledge), and a regulative discourse, which creates order and relations (values and norms). In the pedagogic discourse, the instructional discourse is embedded in the more dominant regulative discourse (Figure 1). In Bernstein's own words:

> Often people in schools and classrooms make a distinction between what they call the transmission of skills and the transmission of values. These are kept apart as if there was a conspiracy to disguise the fact that there is only one discourse (Bernstein, 1996: 46).

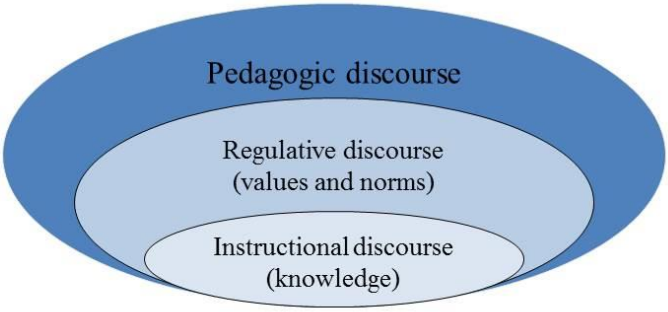


*Figure 1. An illustration of how the instructional discourse is inseparable from and embedded in the regulative discourse.*

The instructional and regulative discourses are used to analyse the messages transmitted by curriculum and assessment in the sub-study presented in article II. Although the two discourses cannot be separated, they are analysed and discussed separately in this thesis in order to make them visible and possible to discuss.

# Visible and invisible pedagogy

When discussing the PE grades of different groups of students, the concept of a visible and invisible pedagogy can help to extend the discussion to *how* the message is transmitted and *which* content is included. Penney also recognises these two dimensions:

> Questions about the knowledge content and forms that are deemed as legitimately integral to physical education are inherently tied to questions of how that knowledge can be communicated, developed and recognised and simultaneously, whose learning needs, interest and abilities will be acknowledged and addressed in the name of physical education" (2013, p. 7).

*Which* content is given priority in PE affects who will benefit. Several scholars have discussed the implications of content for equity, e.g. Hay and MacDonald (2010) and

Evans (2004), who were inspired by Bourdieu's idea of how students' habitus matches the field of PE. Bernstein (2003a) argues that who benefits from the prioritised ways of *how* the message is transmitted also needs to be considered.

Knowledge requirements specify which knowledge is legitimate to grade. The requirements are visible to students and parents alike and, in this thesis, the instructional discourse is shown to be mediated by a visible pedagogy (cf. Bernstein, 2003a, b). The regulative discourse is instead transmitted through an invisible pedagogy, in that no criteria are set at the normative level. Bernstein suggests that a visible pedagogy could make it easier for students who are not familiar with the prioritised knowledge to succeed in education, because the required knowledge is easier to recognise (2003b).

In the third article, and in the discussion chapter, the results are discussed in relation to the macro level of curriculum and agents striving for influence over curriculum and practice. In order to better understand the transformations and the power relations on which production, distribution and reproduction of knowledge rest, Bernstein developed a concept called the pedagogic device, which is helpful when connecting the micro- and macro levels.

## The pedagogic device – part one

The pedagogic device enables the researcher to identify relations between the micro level of pedagogic actions in the classroom and the macro level of steering documents and society. In this thesis, the theory is applied to the pedagogic action of grading in PE. As described by Bernstein (1996, 2003b), the pedagogic device recognises the complexity in the transformations between different fields, where the concept of knowledge is decontextualised and recontextualised into new fields.

The first field, the primary field, is sometimes also called the production field. This is where knowledge is produced (i.e. at universities and research centres). The regulation and distribution of knowledge is ordered by distributive rules that highlight and define "the thinkable and the unthinkable and their entailed practice to different groups" (Bernstein 2003b, p.181). Bernstein (2003b) claims that the relation between power, social groups and form of consciousness is established by the distribution of different orders/ordering of meanings that "can be said to create different knowledges/practices" (Bernstein, 2003b, p. 181). Although Bernstein does not explicitly talk about tacit dimensions of consciousness, the thinkable and the unthinkable, I suggest that PE teachers' perceptions of the thinkable, unthinkable and consciousness could have a tacit dimension.

The second field is the recontextualising field. The knowledge that is generated in universities and research centres is transformed to pedagogised knowledge, for instance in the national curriculum. This transformation is ordered by specific recontextualising rules, where different agents compete for influence over the selection of what is to be transmitted and graded. According to Bernstein (2003b, p. 192), agents and agencies that struggle to influence curriculum and assessment represent two fields: the official recontextualising field (i.e. the state) and the pedagogic recontextualising field (i.e. universities, specialised media and fields that are able to exert influence). The transformation of knowledge always leaves room for influence: "every time a discourse moves from one position to another, there is a space in which ideology can play. No discourse ever moves without ideology at play" (Bernstein, 1996, p. 47). This is crucial to curriculum and pedagogy, because if there is always a discursive gap, there will never be a curriculum or a pedagogic approach beyond ideology (Lilliedahl, 2015). In the thesis, Bernstein's concepts are used to analyse grading in a time period and grading system that are different to those in which the concepts were formulated, which means that the power relations may have changed and new agents may be involved in the struggle for power over the transformations.

The outcome of the transformation of all produced knowledge into a curriculum depends on cultural and political influences. Bernstein (1996) describes two different curriculum models, the competence model and the performance model, which is helpful when identifying who or what influences the transformation. The models are presented below.

## The competence model and the performance model

In the thesis the competence model and the performance model (Bernstein, 1996) are used to problematise how PE teachers' grading is affected when different interests advocate one or the other model. The competence model is based on "'horizontal relationships', the recognition and acceptance of diverse 'abilities', shared needs and interests, and the achievement of personal value and status" (Evans & Penney, 2008, p. 31), where everyone is considered able and differences are seen as complementary. Knowledge has a weak classification and framing and is broadly conceptualised in the competence model. Here, students' knowledge is valued, which means that equity is in focus. Teachers make many different personal assessments in different contexts. Although each assessment is fragmentary, together they can provide an overall picture. Teachers' professional assessments focus on learning and student development and have a formative purpose. Individualised assessments are made at the expense of equivalence, since personal assessments are not particularly comparable (Gipps, 1994). Grades

can serve to motivate students. However, if the criteria are not transparent and the student is not aware of the grounds for assessment, learning can be affected.

The performance model, on the other hand, focuses on "'vertical relationships', differentiation and the creation of 'ability' hierarchies, and the ascription of positional status and value" (Evans & Penney, 2008, p. 31). Here the emphasis is on differences and what is missing. Valued knowledge is conceptualised as limited by the knowledge requirements and has a strong classification and framing in the performance model. The knowledge requirements are suggested to be standardised, measurable and comparable (Sundberg & Wahlström, 2012). Both models exist simultaneously. Which model dominates has changed over time and reflects the different purposes of schooling (Bernstein, 1996, Evans and Penney, 2008). The dominating model affects the transmission of policy to practice, as studied in the thesis and conceptualised in the pedagogic device described in the following section.

## The pedagogic device – part two

When the curriculum is formulated in the recontextualising field, another transformation take place, which is that privileged knowledge in the curriculum is transformed into common or shared classroom knowledge in interactions with students in the secondary field (i.e. primary and secondary school) (Bernstein, 2003b; Singh, 2002). The secondary field is sometimes called the reproduction field, which is where knowledge is reproduced in school. The pedagogic practice is regulated by evaluative roles that "act selectively on contents, the form of transmission and their distribution to different groups of pupils in different contexts" (Bernstein 1996, p. 118). The evaluative rules govern what is recognised as most valued in the school context and, thus, what is assessed and graded.

In the first article, which explores what PE teachers' value when grading in relation to the national grading criteria, Linde's (2012) concept of different arenas[4], is applied. Linde, inspired by Bernstein, is interested in understanding the transformation of intentions in the curriculum into what happens in the classroom. He talks about three different arenas: the formulation arena, the transformation arena and the realisation arena. The formulation arena corresponds to Bernstein's recontextualising field, which is where the steering documents are formulated. The realisation arena corresponds to the secondary field, which is where the intentions in the curriculum are realised at the micro

[4] Linde developed the arenas from Lundgren's work in a curriculum theoretical tradition. Lundgren worked with Bernstein and they wrote together about power relations in education (Bernstein & Lundgren, 1983).

level of the classroom. Linde (2012) also introduces a transformation arena in order to analyse the transformation between the formulation and realisation arenas. The relation between Linde's concept of arenas and Bernstein's concept of fields is illustrated in Figure 2.

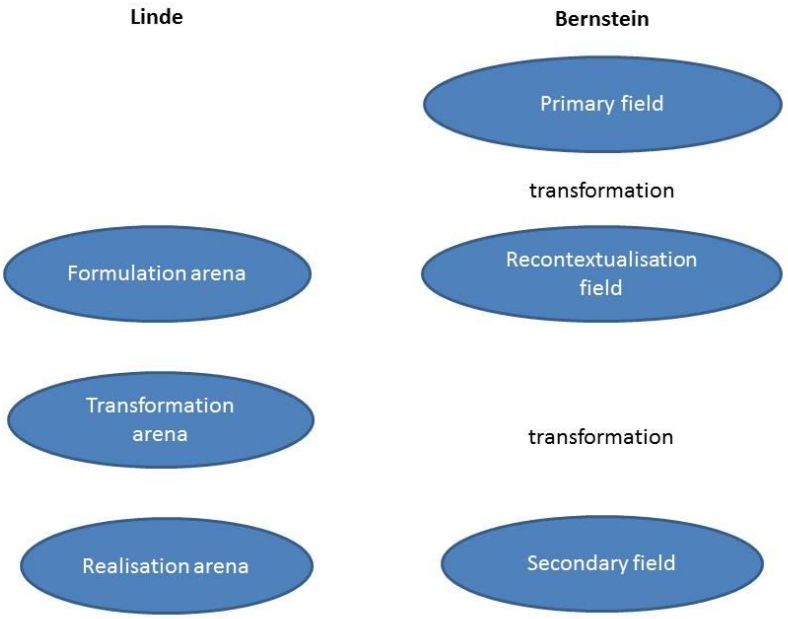


*Figure 2. An illustration of the relations between the arenas proposes by Linde and the fields proposed by Bernstein.*

In the thesis, Bernstein's concept of the pedagogic device is used to analyse teachers' value-setting in relation to curriculum and the different interests that influence curriculum and practice. This enables the discussion to be expanded to include the recontextualising rules, which shifts the focus from the teacher to the structural level.

# 4. Methodology

The first three sub-studies of the thesis are interview studies and the fourth is based on registry data. The former aim to explore, albeit with different foci, what teachers assign value to when grading. The Repertory Grid (RG) interview technique is employed in all the interview studies in order to facilitate the verbalisation of the tacit dimension. The following section outlines the characteristics of such an approach.

## The Repertory Grid technique

The RG technique was developed by George Kelly, an American clinical psychologist (Kelly, 1955), and was originally intended to enable patients undergoing therapy to articulate their own conceptions of their situations with as little involvement from the therapist as possible. The intention was also to enable patients to visualise changes in their perceptions over time. Since then the method has been used in many different fields, such as organisation development, marketing and education. The RG technique has also been employed to investigate the grading criteria in arts and crafts (see e.g. Lindström, 2001). In the field of PE and health it has been used by Rossi and Hooper (2001) to facilitate PE Teacher Education (PETE) students' reflexive constructions of teaching in PE.

The RG technique is based on Kelly's personal construct theory. Here, Kelly (1955) suggests that we humans use systems of constructs which we develop in order to understand and predict the world around us. The constructs are based on our previous knowledge and experiences, which we analyse and try to organise. From these constructs, we generalise, draw conclusions about patterns of cause and effect and adapt our behaviour to them. This testing of assumptions in order to understand the world is something that constantly engages us as humans (Fransella, Bell & Bannister, 2004; Kelly, 1955). For example, if we know that a class is not motivated in PE, we may plan a different lesson content than that for a motivated class. Constructs always have two poles, which are based on opposites and differences. The two poles are not always the result of explicit considerations, but are instead adapted to the concrete situations that

occur. We may characterise a person we meet as parsimonious, which represents one pole. The opposite, often implicit pole, is generous. Kelly (1955) argues that people can interpret everyday phenomena as both similar to and different from other phenomena.

Constructs are not static, but are constantly challenged and can be strengthened, rejected or transformed due to new experiences. If a grading practice is questioned by a student, or if the student does not fit into a teacher's system of constructs, the system can either be adapted or immunised with thoughts like 'the student must have been unusual, this was an exception'. Whether or not people choose to adapt or immunise their constructs depends on several factors: how open they are to new information, how important it is to stick to their convictions that their grading practices are correct, or how important it is for them to have a lot of information about grading. Some of the constructs are more important to maintain or are higher up in the hierarchy, such as those related to a person's fundamental values. The constructs that are high up in the hierarchy are less likely to change, have a greater range of applications and are perceived as more meaningful (Walker & Winter, 2007). Constructs are sometimes described as intuition, gut feelings or perceptions that guide people's actions without necessarily being verbalised (Björklund, 2008b).

Björklund (2008a) suggests using RG to access teachers' tacit knowledge about assessment criteria. I use the RG technique to provide an overall picture of teachers' perceptions of what is important for PE grades. The perceptions are sometimes explicit, transparent and documented in writing, and at other times are intuitive and only found in a teacher's tacit assessment or gut feeling. With the aid of the RG technique, I attempt to identify a teacher's conscious and unconscious constructs of what is important for PE grades. Even though teachers sometimes find it difficult to specify and describe the characteristics of a student they assess, they can usually articulate the differences between one student and another if they know their own students well (cf. Borg, 2008). An advantage with the technique is that it provides a structured way of analysing and presenting data, as expressed by Borell and Brenner (1997):

> The technique can even be said to take note of the qualitative method, the strength, openness to what the respondents actually acknowledge, while trying to get to grips with these methods weaknesses, namely a relatively unwieldy, diffuse form for data presentation and data processing (p. 58).

Borg (2008) points to another advantage, which is that the RG technique facilitates a greater openness to the respondent's own notions and vocabulary. The technique reduces the need to control the respondents using questions based on preconceptions that are different to their own and makes it easier for them to express their opinions in their own words.

## How the interviews are performed

RG interviews are conducted in three steps (Fransella et al., 2004). During the interview, a grid is constructed, as described below. An example of a grid is shown in Figure 3. The grid can be either analysed with an ocular analysis or with a number of software programs. WebGrid5, which is available in free versions from Calgary and Victoria universities, were used in the thesis together with ocular analysis. The results in the first article are analysed with Prin Grid maps, based on principal component analysis. In the first step, generating elements, the teacher selects seven or eight of their own students. The chosen elements (here students) should be homogenous i.e. chosen from the same category or field so that the interviewee is able to compare them in a natural way. The teachers in the study chose the students from a class they had graded and knew well. The elements also have to reflect the variation in the examined category (Borell & Brenner 1997). In the interviews the students represent all the possible grades and are listed in the columns of the grid designated by grade, number and gender (Figure 3).

In the second step, generating constructs, the interviewees articulate their perceptions of what is awarded in PE grades. This is done using triads of students, which reflects Kelly's original use of the technique (Fransella et al., 2004). The names of three of the students are presented to the teacher in various combinations, who is then asked how two of them are similar to and different from the third in the aspects to be graded. The teacher's expressions of the similarities and differences constitute the two poles of the construct. For instance, if a student just wants to work with his/her friend, and this differs from the other two who want to cooperate with everyone, the construct is *just wants to work with his/her friend – cooperate with everyone*. The students are presented in different triads, first with all the possible combinations of grades and then randomly until the teacher can no longer think of any new constructs. Each construct is then inserted into a separate row in the grid (Figure 3).

Finally, in the third step, rating the elements, the teacher is asked to rate all the students discussed on a five-point Likert scale for every construct generated. One of the poles (e.g. just wants to work with his/her friend) represents the scale value 1 and the other pole (e.g. cooperates with everyone) represents the scale value 5. After the rating, every student has a value between 1 and 5 on every construct. An example of a grid in which students represent the columns and the constructs is found in the rows shown in Figure 3.

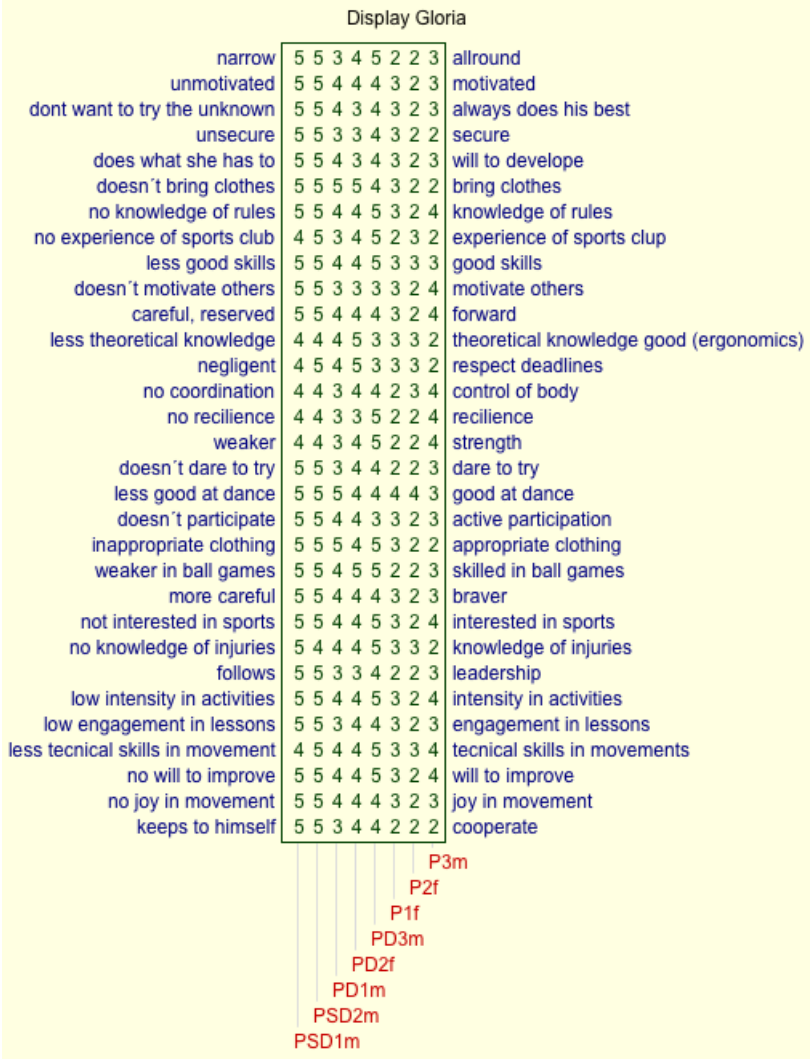


*Figure 3. An example of a grid from 2009. The students in the columns are designated by grade, number and gender. The numbers in the grid show how the teacher has rated the eight students in the constructs presented in the rows.*[5]

[5] The interview in Figure 3 was conducted before the introduction of Lgr11. P represents a pass grade, PD represents pass with distinction and PSD represents pass with special distinction. M stands for male and F for female.

# Selection

The criteria for the selection of respondents in the first three sub-studies were that the teachers have a teaching certificate in PE and have taught and graded students in years eight and nine in nine-year compulsory school. Women and men with varying degrees of experience of grading were selected for the study from local authority schools of different sizes and with different social demographics in two municipalities in central Sweden. Teachers in the selected schools were contacted, informed about the research problem, told that participation was voluntary and that they could withdraw whenever they wished. They were also told that pseudonyms would be used to protect their identity. If the teachers met the criteria they were asked to participate in the study. In the first sub-study, two male and two female PE teachers with different experiences of grading were interviewed. One of the participating teachers was the only PE teacher in a small school, while the other three worked in schools of various sizes with at least one PE colleague, which meant that the opportunity to discuss their grading with colleagues varied.

In the second sub-study, three of the teachers in sub-study one who were still active in the compulsory school setting were contacted and asked to participate in a second interview. Four years had passed since the first interview and by now the 2011 curriculum, with its new knowledge requirements, had been implemented. As it turned out, one of the female PE teachers no longer worked in the compulsory school setting, although two men and one woman did and agreed to be interviewed. Once again they were told that their identities would be protected, that participation was voluntary that they could withdraw whenever they wished and they were also informed about the research problem.

Seven PE teachers – four men and three women – were interviewed in the third sub-study. The teachers represented different sized schools (working alone or with several PE colleagues) in different areas (small village, outskirts or centre of a small and middle-sized town, with different catchment areas). I contacted the selected schools with a view to communicating with the PE teachers. The procedure described in previuos sub-studies was followed.

In the fourth sub-study the aim was to investigate how sociological background factors impact PE grades. National data from SNAE, including data from all the students leaving nine-year compulsory school in 2014 (n= 95317), was used to gain an overall picture. The data included information about each student's gender, PE grade, whether they attended a local authority or independent school, if at least one parent had received an education after upper secondary school and whether they had lived in Sweden since the usual school starting age. The four background variables are dichotomous and the

response variable, PE grades, is classified on a scale from A to F. A classification of NG is given to students who have not been awarded any grade in PE due to an insufficient basis for a decision. The data was provided by SNAE. All personal information and the names of the schools and municipalities had been removed when I received the data in order to protect the students' identities.

## Data collection

In the first sub-study, the four PE teachers were interviewed using the RG technique. The interviews were performed during the spring term of 2009 – two years before the implementation of Lgr11. The teachers had already set the grades for the autumn term, but had not yet set the final grades. The grades of 32 students were discussed. The interviews generated 86 constructs. After the repertory grid interviews, the teachers were asked to rank the importance of their own generated constructs when grading students. A five-point scale was used, where one represented not important and five the most important. The teachers were also asked to make a list of the documentation they used in the grading procedure. The constructs and the documentation were then compared to see whether and how they matched. The four RG interviews took place in private rooms of the respondents' choice and lasted for about 90 minutes each.

In the second sub-study, new interviews were performed in 2013 using the same technique and at the same time of year. This was the first year of Lgr11 in all the year cohorts of nine-year compulsory school. The interviews were recorded and transcribed, which made it possible to compare the teachers' contextualisations of their constructs. The three teachers interviewed in 2009 and 2013 discussed the grading of 32 students, which generated a total of 125 constructs.

Seven PE teachers were interviewed using the RG technique in the third sub-study. The grading of 49 students was discussed and the teachers generated a total of 150 constructs. In addition a semi-structured interview was conducted with the same seven teachers in order to provide rich, detailed answers (Bryman, 2012, p. 470). The teachers were presented with the knowledge requirements for movement stated in Lgr11, one by one. They were also asked to describe how they assessed and graded the knowledge requirements and what kind of problems they faced when assessing and grading each requirement.

In the fourth sub-study the data was collected by Statistics Sweden (SCB) and retrieved from SNAE using information from the Swedish Total Population Register (TPR) and the Swedish Register of Education (UREG). TPR is an excerpt from the population register managed by the Swedish Tax Agency and provides information

about gender, age and year of national registration (SCB, 2016). UREG provides information about education, originating from nine-year compulsory school, upper secondary school and university registers. Data relating to education outside Sweden comes from the population and housing census (FoB) and an annual survey of new arrivals whose educational information is not recorded in UREG (SCB, 2011). In the data retrieved from SNAE, personal identity numbers were missing in 0.4 percent and 2.5 percent of cases relating to parents' education. SNAE chose to substitute the missing cases as having a Swedish background in the former (0.4%) and as no parent having a post-upper secondary school education in the latter (2.5%).

# Analysis

In the first sub-study, the 86 generated constructs were regarded as key thoughts and were coded and categorised using an inductive approach. After becoming familiar with the contextualisation of the constructs, four themes were identified through the data. Prin Grid maps, based on principal component analysis in the WebGrid5 program, were calculated on each teacher's individual constructs to determine how relevant the constructs were for the grade given. The loading of the constructs on the grade relevant component in the Prin Grid maps was then compared with the teacher's own rating of the importance of the constructs in order to explore how well the teachers' ratings matched the construct's relevance for the set grade. WebGrid5 was also employed to calculate how well the constructs matched each other.

The second sub-study had a longitudinal design. The constructs generated in the second round of interviews were categorised into the same themes as the first interviews in order to facilitate a comparison in a content analysis. The distribution of the 125 constructs generated by the teachers on the two separate occasions were compared and interpreted in relation to the curriculum and knowledge requirements.

In the third sub-study, the semi-structured interviews were analysed first. The Swedish transcriptions were read in order to identify the meaning units, label them and group them into categories. The categories were then compared and organised into themes that described the content of the data. Throughout the analysis the data was constantly re-read, re-sorted and re-labelled until an understanding of the content was reached. Thereafter, a summary reflecting the variation was written in English. A new analysis was performed on the RG interviews and when new aspects emerged they were added to the summary.

In sub-study four, the registry data was used to show descriptive data of the relations between PE grades and four background factors. The data was described by frequencies,

percentages and bar charts showing the relative frequencies as percentages for each background factor. In order to study the impact of the background factors on the response variable, namely PE grades, ordinal regression analysis was used (because the measurement scale of PE grades is ordinal). The form of ordinal regression was the proportional odds model, which assumes a constant odds ratio for all cumulative splits of the range of PE grades. The interpretation of the odds ratio is the odds of getting a higher grade than a lower grade by comparing two students belonging to one or the other category of the background factor. The first step was a univariate analysis, where the four background factors were analysed separately. A multivariate analysis was then conducted in which all four background factors were analysed simultaneously, which meant that the effect of one of the background factor was estimated while controlling for the other background factors. In order to determine whether the effects were homogenous across the classification of each factor, stratified analyses were conducted on each factor classification. Specifically, to determine whether the effect of gender, school form and parents' education was homogenous across migration background, a separate multivariate analysis was conducted for each migration background classification.

A new dichotomous variable was created in the light of the results of the regression analysis in order to compare the students with the highest or lowest odds of receiving a high PE grade. The students were classified as highest odds if they had the highest odds in all four background factors in the regression analysis (n=7350) and as lowest odds if they had the lowest odds in all four background factors in the regression analysis (n=2003). The two new categories together included 9,353 students. An ordinal regression analysis was then conducted. All the results were summarised according to the estimated regression coefficients, their standard errors, odds ratios and 95% confidence intervals.

Table 1. Description of the four sub-studies included in the thesis

| | Sub-study I | Sub-study II | Sub-study III | Sub-study IV |
|---|---|---|---|---|
| Sample | 4 PE teachers working in different nine-year compulsory schools | 3 of the PE teachers in sub-study I still working in nine-year compulsory school 4 years later. | 7 PE teachers working in different nine-year compulsory schools | All students leaving Swedish nine-year compulsory school in 2014 (the whole population, n=95317) |
| Data | RG interviews complemented with teachers rating the importance of their own constructs. Data collected in 2009. (Lpo94) | RG interviews. New data collected in 2013. (Lgr11) | RG interviews complemented with semi- structured interviews on how the teachers grade the national knowledge requirements and factors influencing grading. Data collected in 2013. (Lgr11) | Registry data from SNAE (2014) at an individual level of:<br>-PE grades<br>-Gender<br>-Parents' education<br>-Migration background<br>-School provider |
| Analyse | Thematic analysis and Prin Grid maps (principal component analysis) compared with teachers' ratings of constructs | Content analysis | Thematic content analysis | Descriptive statistics and ordinal regression (proportional odds model) |

# Methodological considerations

Validity and reliability can be discussed in terms of trustworthiness (Lincon & Guba, 1985; Guba & Lincon, 1994; Bryman, 2012; Hassmen & Hassmen, 2008). Trustworthiness can be divided into credibility, transferability, dependability and confirmability (Bryman 2012). In the interview studies, efforts were made to ensure that good research practices were adhered to. Respondent validation was performed during the RG interviews, which gave the teachers the opportunity to reflect on each generated construct in their rating of the students. A follow-up meeting with each teacher was held to confirm the identification of the themes in the first sub-study, which enabled me to discuss the constructs and the categories, correct any errors and discuss the interpretations. The other interviews were recorded and transcribed, which made it possible to identify and compare the teachers' contextualisations of their constructs and to continuously check and re-check them. As the RG technique was new to me as an interviewer, two pilot interviews were conducted so that I could familiarise myself with the technique and identify any weaknesses. One adaption was that the names of the students were written

on pieces of paper and presented to the teacher three at a time, due to the difficulty of remembering all the names of the students under discussion.

Confirmability means the degree to which the outcomes can be confirmed by others. This was approached differently in the sub-studies. I used an audit strategy in the thematic analysis described in article I. A peer researcher categorised the same constructs and ended up with the same categories, but with two constructs interpreted in a different way and placed in a different category. The two differing constructs were categorised after a follow-up meeting, where the respondents who had generated the constructs could explain what they meant. The teachers were also invited to comment on the categorisation of all the constructs. In articles II and III, the teachers' reflections when generating the constructs were recorded and transcribed. During the categorisation process, the transcriptions were read and re-read several times in an attempt to ensure that the constructs were interpreted as the teachers intended. The semi-structured interviews in article III were also recorded and transcribed.

Reflecting on my own knowledge and experiences of the study object has been important, because they could affect the availability of the field, the analysis and also the results. Being a PE teacher myself and having considerable experience of teaching and grading has influenced the direction of the study, in that it addresses the questions I have come across in the profession. It also means that I am part of the culture under investigation. This could be advantageous, in that I am familiar with the context, but is disadvantageous if I make interpretations based on my own preconceptions instead of the factual data. As Hastie and Hay (2011) remind us, it is important to look at and interpret the data as a researcher, not as a PE teacher. The data was reflectively analysed and constantly checked and re-checked during the entire process in order to ensure and improve confirmability.

The RG technique has both advantages and disadvantages. One disadvantage is that it deals with individual perceptions, which means that the results are difficult to transfer. Transferability is also difficult to achieve in small samples, which tend to be oriented to the contextual uniqueness (Bryman, 2012). Instead, a thick description can provide others with a database for deciding the possible transferability to other contexts (Lincon & Guba, 1985). In an attempt to be as transparent as possible, and provide a number of thick descriptions, tables with generated constructs are included in the articles. The tables also help the reader to evaluate the analysis and thereby strengthen the dependability. Peer auditing is uncommon in qualitative research (Bryman, 2012), although constructive peer assessments have been performed by reviewers in different journals. The closest thing to auditing in this thesis is the thematic analysis in article I, where a peer researcher made the same analysis as that described above. An ambition has been

to provide a thick description of the methodology, especially the RG technique which many are unfamiliar with, in order to improve dependability.

The fourth sub-study can also be discussed in terms of validity and reliability. One advantage with a secondary analysis of official statistics is that it can offer new interpretations. A disadvantage can be that key variables are missing (Bryman, 2012, p. 315). When performing a multivariate ordinal regression analysis, the effect of one of the background factors is estimated while controlling for the other background factors. However it should be noted that they do not control for any other possible factors that co-vary with the four background factors analysed. Unfortunately, the variable *students' sporting experience*, which has shown to impact PE grades (Eriksson et al., 2005), is not available, because no registers are kept with that information. However, the register from SNAE provides data for the population of students leaving nine-year compulsory school in 2014 (n=95317). The data is collected by SCB from different registers (TPR and UREG) and not individually reported, with the exception of cases of foreign education in UREG. This is an advantage, in that flaws are less likely. A score test of parallel lines was conducted to test the proportional odds assumption. Although the tests most often turned out to be statistically significant, we accepted the proportional odds model in order to keep the models simple and because the large sample size meant that even small departures from the null hypotheses would become significant.

# Ethical considerations

The ethical requirements for information, approval, confidentiality and how to use the information proposed by the Swedish Research Council (2016-02-10) have been adhered to. The requirement for information means that the researcher should inform the interviewee about the aim of the study. In this case, the teachers were informed about the aim in my first telephone conversation with them, during which I presented myself, my background, the aim and the procedures that would be followed. The requirement for approval was dealt with at the same time. Here, the teachers were asked to participate if they fulfilled the criteria, to which they then responded. I also gave detailed information about the possibility to withdraw from the study at any time without having to explain why. The interviewees were further informed that their responses would remain confidential and that pseudonyms would be used in the written reports. Information about individuals should only be used for research purposes and it was made clear that this applied to all the material that was collected, the analysis of grids, recordings and transcriptions. All the data is kept securely under lock and key at the University of Gävle and is not accessible to any unauthorised person.

None of the studies meet the conditions for an ethical review, as suggested by the Ethical Review Act (SFS, 2003:460). In the first three articles no registers are kept. The data in article IV is provided by SNAE, after randomly decoding for personal information, school and municipality (randomly numbered) to protect students' anonymity. No key exists that can make identification possible. Therefore, it cannot be considered as personal data and does not need to be reported to the Swedish Data Protection Authority. Neither is the data what the Personal Data Act (SFS, 1998:204) considers sensitive, because it does not reveal information about race, ethnicity, political opinion, religious or philosophical beliefs, trade union membership, health or sexual tendencies. Consequently, the study is not subject to the Ethical Review Act.

# 5. Summary of articles

In the following sections the results in the four articles, focusing on one research area each, are summarised. Table 2 is presented thereafter to give an overview of the results.

## Article I: Exploring PE teachers' 'gut feelings': An attempt to verbalise and discuss teachers' internalised grading criteria

*Background and aim*: PE teachers occasionally refer to "gut feelings" when setting grades. Previous research shows that their grading is sometimes inconsistent with the official grading criteria. Against this background, the aims are to explore what PE teachers consider important when grading and how their transformation of the official criteria into grading practice can be understood.

*Theoretical framework and methodology*: Linde's concepts of the formulation arena, the transformation arena and the realisation arena are used to focus on the transformation of curriculum into grading practice. The RG interview technique is employed to facilitate access to seldom verbalised criteria and the values teachers give to them. Four PE teachers were interviewed in 2009 before the implementation of the new curriculum, Lgr11 and asked to rate the importance of each construct (criteria) for the PE grades.

*Results and conclusions*: Four themes are identified as a result of the teachers' constructs: *motivation, knowledge and skills, confidence* and *interaction with others* in a decreasing order of frequency. The theme *knowledge and skills* is the closest to the official criteria, although the constructs in the theme do not cover all the official learning outcomes. The missing constructs refer to activities that are performed outside the gymnasium, e.g. swimming, orienteering, winter sports, outdoor life and what can be considered as theoretical knowledge. Aspects relating to knowledge and skills, but which are not included in the official criteria, such as sporting experience and physical qualities, are also mentioned in the teachers' constructs. The official grading criteria focus on subject knowledge. The remaining three themes, *motivation, confidence* and *interaction with others*, instead focus on students' characteristics. However, when ana-

lysing the teachers' constructs in individual Prin Grid maps, all four themes are relevant for the grades set. The teachers sometimes have difficulties predicting which constructs are relevant to the grades set, and the constructs they consider important are not always reflected in the grade.

The verbalised criteria reveal that teachers sometimes use grades to encourage positive student behaviour, to help them deal with the classroom situation and to facilitate students' learning. School facilities and access to learning environments other than the gymnasium also seem to influence teachers' transformation of the official grading criteria to grading practice. The results illustrate that the significance of individual constructs is sometimes hidden, even from the teachers. In order to become cognisant of and develop their grading, teachers need to verbalise their individual grading criteria. The RG technique is a way of making individual constructs in the local context visible and possible to discuss. The results provide discussion points about why the teachers grade the way they do.

## Article II: Swedish PE teachers struggle with assessment in a criterion-referenced grading system

*Background and aim*: In many subjects grading has a long tradition of being influenced by students' characteristics. Sweden has had a goal and criterion referenced grading system since the middle of the 1990s. The Swedish knowledge requirements for PE, introduced in the 2011 curriculum, are expected to strengthen the governance of knowledge as the sole basis for grades. The aim of the article is to explore the impact of the new reform and the stated knowledge requirements for what teachers value when setting grades in PE. The intention is also to discuss which non-knowledge related aspects (if any) teachers continue to look for and why these seem resilient to the reform.

*Theoretical framework and methodology*: The interrelated message systems of curriculum, pedagogy and assessment, as proposed by Bernstein (2003a), are used to discuss what the teachers value-set when grading. Three of the teachers who participated in the first sub-study conducted in 2009 are revisited in 2013. New interviews using the RG technique are conducted and the constructs the teachers generate this time around are compared with those from the first round of interviews in 2009.

*Results and conclusions*: A noticeable change can be detected in what the teachers consider important for PE grades. The teachers generate almost twice as many knowledge constructs and half as many motivation constructs. Constructs based on confidence and social skills are almost eliminated. After the implementation of specified knowledge

requirements for different grades, the teachers talk about theoretical knowledge more often and in new ways. Planning and reflection on the activities are now in focus, in line with the knowledge requirements. A health discourse is also visible. A discourse equating grading of knowledge with best practice seems to have made an impact on the teachers' grading practices. The knowledge requirements appear to direct the teachers' attention to knowledge and skills and away from students' behavioural characteristics.

However, the results also indicate that clear knowledge requirements alone do not lead to teachers ceasing to grade motivation. In order to better understand the persistence of non-knowledge aspects as a basis for grading, Bernstein's (2003a) concept of the interrelated message systems of curriculum, pedagogy assessment and the alignment between them is employed. Curriculum transmits a message about the importance of the values and norms that teachers are expected to teach. However, as values and norms are not included in the knowledge requirements, they are not officially graded. This seems to be problematic for the teachers in the study. Bernstein (1996) emphasises that the transmission of skills and the transmission of values cannot be separated. According to the knowledge requirements, knowledge is more important than values and norms, which sends a different message than the overall curriculum, which emphasises that they are all of equal importance. Teachers sometimes seem to compensate for this inconsistency by creating an alignment between what is communicated by the curriculum and what is graded.

The results point to teachers' struggles with the differences between the specific criterion referenced grading and the more process oriented non-criterion referenced learning. The results also raise the question about how to achieve valid grades and at the same time acknowledge the need to signal the importance of values and norms in order to accomplish all the goals stated in the curriculum.

## Article III: Swedish PE Teachers' Understandings of Legitimate Movement in a Criterion-referenced Grading System

*Background and aim*: Earlier research shows that, in PE, the subject content is related to sport- and activity discourses at the practical level, while the normative level proposes a broader and more inclusive concept of movement. Grading should be based on movement qualities, rather than results in competitive sports. Against this background, the aim is to investigate which kinds of movement PE teachers consider legitimate to grade in a criterion referenced grading system and which factors influence their grading.

*Theoretical framework and methodology*: Teachers are regarded as active mediators in the transformation described in Bernstein's (1996) concept of the pedagogic device, where for instance the national knowledge requirements for movement are transformed into the kinds of movement that are valued in teachers' grading practices. Semi-structured interviews were conducted with seven PE teachers who also were interviewed using the RG technique.

*Results and conclusions*: The teachers in the study often preserve the established discourses rather than use the grading system to challenge traditional knowledge boundaries of movement. The results confirm earlier results, namely that a sport and activity discourse frames teachers' interpretations of what constitutes legitimate movement. Influences from sport federations and competitive sport can be detected in the teachers' interpretations of the knowledge requirements, which results in students' speed, strength, techniques and tactics being valued when grading.

The teachers' difficulties in communicating movement qualities and their pedagogical concerns about how to motivate all students to learn help us to understand their grading practices. The results reveal a tension between demands for transparency and the grading of movement qualities. In a high stakes grading system, teachers are expected to explain to the student what is required for a specific grade. On the other hand, qualitative movement requirements can be difficult for teachers to articulate, especially if the knowledge requirements are vaguely formulated, such as *to some extent/relatively well/well* (SNAE 2011). This tension occasionally leads to quantitative measurements of movements being preferred in order to facilitate the communication with students and parents. The knowledge requirement, *to be able to adapt movements to activities and context* (SNAE, 2011), is sometimes interpreted as results in competitive sports or the number of exercises performed.

Teachers are also concerned about students' unequal possibilities to develop, practice and display the knowledge requirements due to socioeconomic factors, group dynamics and movement experience. Occasionally, teachers compensate for students' different possibilities by adapting the knowledge requirements, thereby making them possible for the student to reach. This can be interpreted as teachers reacting to tensions between demands for equivalent grading and pedagogic concerns to encourage everyone to learn according to their own abilities.

# Article IV: Who gains? Sociological parameters for obtaining high grades in physical education

*Background and aim*: Equity in the Swedish school system has decreased in the last two decades. The aim is to identify and analyse sociocultural factors that contribute to the success of students in Swedish PE. The focus is specifically on how gender, parents' education, migration background and school provider impact PE grades in the context of school marketisation and an increasing focus on results and grades.

*Theoretical framework and methodology*: Bernstein's concept of the pedagogic device is applied to discuss socio-political influences on the school system and equity. Registry data relating to all students leaving nine-year compulsory school in Sweden in 2014 (n=95317) is accessed from SNAE and ordinal regression analyses are performed to estimate the relation between sociocultural factors and PE grades.

*Results and conclusions*: Descriptive data shows that boys, students having at least one parent with a post-upper secondary school education, students living in Sweden since starting school and students attending independent schools have higher grades in PE than those who do not. All four background factors have a significant effect of higher grades in ordinal regression analyses, both separately and when controlling for the other factors. In the multivariate calculations, the odds ratio for receiving a higher grade in PE is higher if a student has lived in Sweden since starting school, has at least one parent with a post-upper secondary school education, attends an independent school and is a boy, in decreasing order.

In order to explore how background factors work together, a new dichotomous variable is created in the light of the multivariate calculations. Students are classified as highest odds if they are boys who have attended independent schools, have at least one parent with a post-upper secondary school education and have lived in Sweden during their nine-year compulsory schooling (n=7350). The students are classified as lowest odds if they are girls who have attended local authority schools, have no parent with a post-upper secondary school education and immigrated to Sweden after the ordinary school start (n=2003). The highest grades, A and B, are awarded to 42 per cent of students in the group highest odds, but only to 7 per cent in the group lowest odds. The group with highest odds has almost six times higher odds of getting a higher grade in PE than that with the lowest odds.

The strong impact of parents' education on PE grades in the results reflects the inequalities found in society. High education and socioeconomic status is associated with a higher activity level (Elofsson, et al., 2014; Engström, 2008) and more frequent and active membership in sport clubs (Elofsson et al., 2014; Norberg, 2016). As a subject,

Swedish PE has changed from one of doing, using the body for motor learning and promoting sport skills to one of knowing and promoting reflective action and insights into health issues (Annerstedt 2005). This is sometimes regarded as an academisation of the subject, in that the ability to analyse and express knowledge verbally or in written texts is a requirement. Statistically girls, students with parents with a post-upper secondary school education and students born in Sweden are associated with higher grades in theoretical subjects and are expected to benefit from such academisation. The results in this study show that these students have higher grades, but that girls have a lower grade level than boys. The persistent inequality revealed in the study indicates that official compensatory equity policies have failed to compensate for influences of power structures framed by gender, class, ethnicity and culture.

# Summary overview

Table 2. An overview of the results in the four articles

| | |
|---|---|
| Art. I | -Four themes are found in teachers' grading, only one of which addresses knowledge.<br>-The teachers' constructs reflect the classroom situation and restricting factors.<br>-Teachers' difficulties in predicting the relevance of their own constructs for the grades given suggest a tacit dimension. Matching constructs can be a possible explanation of why constructs become relevant for the grade given even though they are not considered important by the teacher. |
| Art. II | -The implementation of official knowledge requirements seem to impact teachers' grading practices and improve alignment.<br>-Motivation prevails as a ground for assigning grades.<br>-Curriculum goals contain knowledge, values and norms, but the knowledge requirements prescribe knowledge as the only ground for grades. The results suggest that this inconsistency is problematic for the teachers. One interpretation is that by grading motivation the teachers compensate for inconsistency in what Bernstein calls the interrelated message systems of curriculum, pedagogy and assessment. |
| Art. III | -Teachers seem to preserve existing structures for legitimate movement rather than challenge them.<br>-Teachers' interpretations of legitimate movement are influenced by an activity discourse and competitive sport. As a result, fitness is given value in the grades.<br>-Teachers mention dilemmas that could help to understand their practices, for instance:<br>-Tension between demands for transparency and difficulties in explaining the qualitative requirements for movement in words.<br>-Tension between demands for equivalent grades and pedagogic concerns to encourage students with different background and possibilities to learn according to their own abilities. |
| Art. IV | -Migration background, parents' education, school provider and gender impact PE grades, in decreasing order.<br>-When combining the factors with the highest odds for higher PE grades, the odds for getting a higher grade are almost six times higher than when combining the factors with the lowest odds. The result points to the importance of an intersectional approach in research.<br>-The influences of different background factors signal which knowledge is valued and facilitates a discussion about how PE can challenge the existing structures and enable more students to be included and feel able in the PE context. |

# 6. Discussion

The overall aim of the thesis is to explore grading in PE. The results of the sub-studies are discussed in this section, beginning with the first focus, which is to explore what teachers value in their grading. This generates three research areas, which are presented in three separate articles. The results have implications for equity, which leads to the second focus, namely to identify and analyse the sociocultural factors that contribute to the success of students in Swedish PE assessment activities. This is attended to in article IV and provides a backdrop to discussions about the results presented in articles I-III. I begin the discussion by reflecting on some of the strengths and weaknesses of the thesis.

## Critical considerations

Some critical considerations of the thesis need to be expressed. The first is that the aim to investigate what teachers value when grading includes the exploration of a tacit dimension. I am aware that a tacit dimension can never be fully understood, precisely because it is tacit, although the more we know, the better able we will be to improve equivalency and discuss the criteria.

In my view, the use of the RG technique is a strength, in that it facilitates a deeper insight into teachers' personal understandings of grading. The technique makes it possible to investigate teachers' personal perceptions of what they consider important for PE grades and to analyse the internal validity of the constructs. On the other hand, the RG technique can also be a weakness. As it is a qualitative method, the interviews are conducted with small samples and no generalisations can be made. Nevertheless, by adapting Bernstein's conceptual framework, it would be possible to contextualise new situations and contexts in the framework provided.

Bernstein's conceptual framework provides possibilities to connect the micro- to the macro level. However, it is also possible that theoretical frames may prevent us from seeing alternative interpretations, as expressed by Gard:

> …working within the theoretical framework proposed by an important educational theorist like Basil Bernstein […] can generate important insights. But a largely ignored question of

> at least equal importance is what this theoretical framework prevents us from seeing. All theories conceal and reveal, producing knowledge that we decide to celebrate while consigning other knowledge to what the sociologist Zygmunt Bauman characterised as an intellectual 'waste' (Gard, 2015, p. 115).

The results of the thesis are discussed in the light of Bernstein's conceptual framework. This enables me to detect patterns that might remain hidden without the concepts. At the same time, it may prevent me from seeing other possible interpretations that have not been in focus in this thesis.

I will begin the discussion of the results by attending to the equity aspects dealt with in article IV and then continue with the equity implications of the findings in the first three articles. Thereafter, the results reported in the first three articles are discussed by identifying the agents that become visible in the transformation of the curriculum in the recontextualising field into teachers' grading practices in the secondary field. Finally, the values and purposes in the discourses that become visible in the results are conceptualised against the competence and performance models suggested by Bernstein. This enables a contextualisation of the problematic situations that the teachers draw attention to in the three first sub-studies.

# Equity in PE

The results presented in article III point to PE teachers trying to deal with students' different possibilities to meet the knowledge requirements in PE in a goal referenced grading system. Equity is an important goal in the Swedish school system. However, there are now signs that this is decreasing (SNAE, 2006, 2014). According to OECD, "the highest performing education systems across OECD countries are those that combine quality with equity" (2009, p. 9). Education in Swedish schools aims to compensate for inequalities and promote students' development by adapting the teaching to students' circumstances and backgrounds (SNAE, 2014).

However, the results presented in article IV show that the odds of receiving a higher grade in PE are greater if you are a boy, have lived in Sweden since starting school, attend an independent school and have at least one parent with a post-upper secondary school education. The teachers featured in article III mention sporting, socioeconomic and migration backgrounds as making it difficult for some students to meet the knowledge requirements. The last two mentioned background factors – socioeconomic and migration background – are in line with those with the highest odds ratio of receiving a higher PE grade, namely parents' education and migration background, as reported in article IV. Unfortunately, no data exists for sporting experience, although in a national evaluation of PE, sporting experience, as reported by students had the highest odds

ratio of the background factors (Eriksson et al., 2005). Boys in the last years of nine-year compulsory school, children with a Swedish background and children with higher parental education tend to be more active in sport federations (Elofsson et al., 2014; Norberg, 2016). The differences in participation in sports federations between children with a foreign or Swedish background are small amongst the boys but larger amongst the girls (Norberg, 2016). This is in line with the results reported in article IV, where the differences in the chances between boys and girls receiving higher grades in PE are larger in the group moving to Sweden after the ordinary school start. This indicates that girls with a foreign background are disadvantaged in physical activities, both at school and in their leisure time.

The importance of sporting experience for PE grades is connected to the PE lesson content and the sport discourse, both of which are reflected in the teachers' interpretations of legitimate movement reported on in article III. Only including the knowledge that is stated in the knowledge requirement benefits students who are familiar with the content. Swimming and outdoor activities are recognised as those that students born and brought up in Sweden are familiar with (Eriksson et al., 2005). A health discourse in PE could also affect equity. PE has sometimes been considered part of the solution to public health problems and child obesity (Kirk, 2010; Quennerstedt, 2006). Evans (2014) discusses PE´s responsibility for health and obesity issues to reproduce existing inequalities, since health is often interpreted as being equal to slim and fit. Thereby, a fat body is associated with poor ability in PE and poor health, thus reproducing existing social hierarchies, "not despite our profession's best intentions to address the wellbeing of children, but because of them!" (Evans, 2014, p. 321). When health is interpreted as being slim and fit, groups with a higher socioeconomic status can be advantaged, especially as they are associated with a healthier lifestyle and higher activity levels (cf. Engström, 2008).

According to the results presented in the first article, teachers' expectations based on experience may impact their grading, for example by taking factors that are easy to detect into consideration if they match the desired abilities according to their own experience. There is a risk that this could contribute to the reproduction of existing inequalities, particularly as the results reported in article IV suggest that experience tells teachers that boys, students starting their school life in Sweden, students attending an independent school and students with at least one parent with a post-upper secondary school education perform better in PE. Teachers' expectations can for example make it more difficult to detect students who do not fit into the experience-based pattern, thereby reducing the possibilities to challenge existing structures. Unless expectations based on experience are made visible and reflected on, they may work "behind the backs" of

teachers and students in ways that reproduce existing social structures (Biesta 2015, p. 77).

Bernstein's conceptual framework enables us to investigate *how* the message is communicated by using the concept of visible/invisible pedagogy, where visible pedagogy is embedded in invisible pedagogy (Bernstein, 2003b, p. 84). Both forms of pedagogy carry cultural bias, including that of social class. Visible pedagogy, with its clear criteria, can help to lessen the differences in educational outcome due to the home situation (Bernstein, 2003b). Bernstein (2003b, p. 79) suggests three possible ways of facilitating this: a supportive preschool to compensate for different pedagogic contexts in the home, a relaxing of the frames for how quickly a child has to learn things in a certain order and a relaxation of the frames for communication between the school classroom and the community(ies) that schools draw on.

Sweden has a well-developed preschool system and schools constantly work to improve communications with students' families and the surrounding society. The second suggestion, relaxed frames for how quickly thing should be learned, has crucial implications for the economy, teacher training and school management (Bernstein, 2003b, p. 79). Sweden has instead gone in the opposite direction by prescribing which knowledge should be learned at different ages. The stronger framing of the learning pace could counteract equity due to the different starting positions students have from home (Bernstein, 2003b). Officially, Sweden has turned to what Bernstein calls a repair system to cope with students who have failed to live up to expectations. Another strategy suggested by Bernstein is to reduce the quality or quantity of the content, or both (2003b, p.74). As discussed in the third article, this strategy is employed by some teachers, for example by grading a lower quality of movement as a pass. An invisible pedagogy can be more difficult to understand for students who are not familiar with the curriculum's descriptions of how to think, talk or move, perhaps due to their family backgrounds. As argued in article II, the instructional discourse employs a visible pedagogy in the sense that knowledge requirements are transparent to all. The regulative discourse is instead expected to be transmitted through an invisible pedagogy, given that no criteria are set. When it comes to values, the aim is that: "The school should actively and consciously influence and stimulate pupils into embracing the common values of our society, and their expression in practical daily action" (SNAE, 2011, p. 14). Common values are to be in "accordance to ethics borne by Christian traditions and Western humanism" (SNAE, 2011, p. 9). Sweden has changed from a relatively monocultural country into one that is multicultural. In today's Swedish society, as common values may no longer be all that common, a question to ask is whether an invisible pedagogy will enable students to understand the common values and norms if they are not accustomed to them at home. Some students obviously recognise the regulative discourse as a basis for grades.

When the students in Sivenbring's thesis (2016) found the knowledge requirements difficult to understand, they employed the strategy of displaying generic knowledge about being a "good student" to receive higher grades. PE students also recognise the regulative discourse as a strong ground for grading (Redelius and Hay, 2009). However, if students are unable to recognise what is embedded in the regulative discourse this could affect equity. My interpretation of the results in articles I and II is that the regulative discourse impacts grading. Research also shows that teachers tend to focus more on the regulative discourse when grading students with lower socioeconomic backgrounds (Klapp Lekholm & Cliffordson, 2008). If students are already familiar with the prevalent values and norms, there seems to be less need to signal their values by using the message system of grading.

## Agents in the transformation of policy into practice

In the first three articles I investigate how the curriculum transforms into teachers' grading practices. In the transformation of the recontextualising field (curriculum) into the secondary field (teachers grading practices), different agents strive to influence the evaluative rules. The agents that I discern as influencing this transformation in the first three articles are described below.

Teachers are obvious agents in that they are able to adapt new reforms or resist them. To use Penney's words: "PE teachers are active mediators of both progressive and regressive policy discourses" (2013, p. 8). As shown in articles I and III, teachers' beliefs, which are in turn influenced by the past, present and future, affect their interpretations of the knowledge requirements and relate to the additional criteria they use when grading. Biesta, Priestly and Robinson (2015) find that teachers' beliefs are strongly oriented against the here and now, i.e. the present. When teachers express goals for the future they are often short term. This is also the case for the PE teachers described in the first three articles who are engaged with solving the problems in the classroom situation that affect their grading. The values that are assessed in a PE class can, in part, be attributed to the official criteria in the recontextualisation field. However, teachers' own values, beliefs and expectations do not only influence their interpretations, but may also provide additional and alternative contributions to the dynamics of the field (Hay & Penney, 2013, p. 45).

Similar to the findings in Scotland (McPhail, 2007), there is a tension between the recontextualisation field and the primary field, which McPhail suggests is promoted by teachers who are not involved in the formulation of the curriculum, but are nevertheless expected to reproduce the knowledge contained in the externally produced curriculum.

The introduction of specified national knowledge requirements can be considered to increase the state's influence over the evaluative rules in the Swedish context, thereby limiting the influence of the teachers. This can be regarded as an attempt to replace professional judgement with uniform criteria. In articles II and III it is shown that if the knowledge requirements are well formulated and clear, they are easier for teachers to follow.

Students are another group of agents in the transformation of policy into practice. As discussed in article III, by supporting or resisting the prevalent practice, students can influence the workings of the system and induce changes or adaptions. A sport and activity discourse influences the content of the lessons, as shown in article III. Many students have been, or still are, members of different sport federations (Thedin Jakobsson, et al., 2012) and therefore endorse competitive sport. They have experience of and are able to perform well. Other activities can cause resistance, and there are examples in article III of students refusing to participate in activities if they are unsure about being able to master them. There are also students who are not comfortable doing PE and experience marginalisation and exclusion in the PE context (Eriksson et al., 2005).

The school management, students and parents can also be considered agents in a decentralised school system with a free school choice, such as that in Sweden. Mickwitz suggests that teachers are pressured by demands for higher grades from these three agents (Mickwitz, 2011). However, the teachers taking part in the third sub-study, all working in local authority schools, only mention demands for higher grades from parents and students.

Different dilemmas emerge in the teachers' grading decisions that reflect the interests of the different agents at the macro level. Biesta et al. (2015) claim that "access to wider discourses about teaching and education would provide teachers with a perspective on the beliefs they and their colleagues hold, and would provide a horizon against which such beliefs can be evaluated" (p. 638). In line with this reasoning, I will attempt to highlight the influences that are visible in the curriculum as a horizon against which the dilemmas that teachers experience when grading can be evaluated. Hopefully, this contextualisation will enable teachers to identify dilemmas as part of the struggle between wider discourses and become engaged in discussions about privileged discourses and thereby be in a better position to challenge them. The performance model and the competence model proposed by Bernstein (1996) help us to make connections between the micro- and macro levels, as described in the next section.

## The relations between the micro- and macro levels

The analytical aim of the thesis is to analyse and problematise teachers' grading practices in relation to the different agents and agencies in society vying for influence over the content of curriculum, as described by Bernstein (2003b). The pedagogic device offers a way of studying the power relations in the construction of teachers' grading practices. The competence model and the performance model (Bernstein, 1996) are applied in order to make the influences on teachers' practices from the macro level visible. These two models are considered to be the outcome of the transformation of produced knowledge in the primary field into a curriculum in the recontextualising field. Rather than give a full picture of the complex web of powers that is reflected and reproduced in the Swedish curriculum, my intention here is to outline some of the influences that have emerged from the results of the sub-studies in order to illuminate how assessment activities are impacted by social and cultural contexts.

According to my interpretation, the discourses that are visible in the results are a neoliberal discourse, a pedagogic discourse, a sport discourse and a health discourse, all of which represent different values and purposes. How these values and purposes can be aligned with the competence and performance models proposed by Bernstein (1996) is described below. However, I acknowledge that the overall picture of two models is more complex and has many variations (Bernstein 1996). The two models can be more or less dominant at different times and in different countries. They can also be used differently for different students at the same school or in the same class. In order to make the two models visible in the discussion about the influences on curriculum, the distinctions are retained and adapted to a Swedish context.

In this analysis, the values and purposes in the competence model are considered to align with a pedagogic discourse that dominates the overall curriculum. This means that students' learning and development are in focus and that the pedagogy is to be adapted to each student's individual needs in order to give everyone the possibility to reach the goals regardless of their different background factors (SNAE, 2011). The individual's development of knowledge, values and norms is equally highlighted. As Bernstein emphasises, in the pedagogic discourse the instructional discourse is embedded in the regulative. In the knowledge requirements, the competence model can be detected in the qualitative descriptions of the different grade levels, thereby leaving room for teachers' professional assessment and giving knowledge a weak classification. PE is sometimes seen as part of the solution to public health problems and child obesity (Quennerstedt, 2006). As a result, a health discourse has been given increased influence over the subject at a normative level in the steering documents. This is visible in the goals stated in the PE syllabus relating to knowledge about "how they can safeguard their health

throughout their lives" and to adopt a lifelong physical active lifestyle (SNAE, 2011, p. 50). These values and purposes can be regarded as aligning with the competence model in the sense that they aim at a student's' personal development into a physically active person with a healthy lifestyle and allow for different ways of accomplishing this.

The performance model is visible in the neoliberal discourse, which advocates a measuring culture with high demands for transparency and accountability. Bernstein describes the twentieth century as a time when education has strong links to the economic field as well as to the state and is "regulated by a State operating an anti-collectivist, market-based ideology" (Bernstein, 2003b, p. 157). Reforms to make education more effective have been implemented in an attempt to achieve quality by measurement, competition and freedom of choice. Grades, besides being used for selection of students for higher education, also have the function of evaluating schools. The new function of grades as an instrument for evaluating schools puts the focus on equivalence in grading. The Swedish knowledge requirements can also be considered to represent a performance model, where students become aware of what is missing to reach the next level of knowledge.

In the neoliberal influenced performance model, grades are intended to be the same, regardless of which school the student attends or which teacher the student has, thus suggesting a strong classification of knowledge. The performance model does not take into account students' different starting positions, in that everyone is expected to be measured by the same standards. Historically, sport federations can be considered agents in the PE context (cf. Lundvall &Meckbach, 2008; Olofsson 2007) in the sense that PE content is inspired by sport federations and competitive sports (Ekberg, 2009; Eriksson et al., 2005). Although the content is still inspired by a sport discourse, its influence on steering documents has declined and results in competitive sports are no longer requested. The influence of competitive sports on the content of PE can be regarded to align with a neoliberal influenced performance model, in the sense that hierarchies are created and also that the results of competitions are quantitative, well-defined and measured in the same way in every competition in order to be comparable.

My interpretation is that the discourses that are described as influencing curriculum and grading can be regarded as ideologies at play. I suggest that the assessment practices reported in this thesis are, to use Lingard, Mills and Hayes' (2006) words, "not arbitrary but reflect struggles over the purposes of schooling, as much as the nature of assessment practices themselves" (p. 87). In the next section I relate some of the dilemmas raised in the sub-studies to the competence and performance models, with the intention of illuminating how teachers' dilemmas relate to structural dilemmas.

## Contextualising three dilemmas in the teachers' grading practices

The competence and the performance models exist in parallel, although they have separate purposes and values that sometimes coincide and are occasionally in conflict and cause dilemmas for teachers when grading. Examples are provided of how the two parallel models can contextualise the dilemmas that the teachers in the study experienced. I start with the dilemma of qualitative criteria, indicating a competence model, and the demands for accountability from parents and students, indicating a neoliberal influenced performance model, as described in article III. The qualitative criteria for movement depend on a teacher's subjective professional assessment of movement qualities that are sometimes difficult to describe in words (see also Nyberg & Larsson, 2014). The assessment of movement qualities can be easy for parents and students to question, because they often expect the teacher to describe in words exactly what is missing in order to reach the next grade level. The teachers reported on in article III tend to act in different ways depending on the situation they are in. High demands for accountability can be challenging, for instance when teachers are questioned by parents and students and forced to defend assessments verbally, especially when the requirements are described in vague terms, such as *to some extent*, *relatively well* and *well*. In some cases these demands for accountability make the teachers in article III measure movement quantitatively, for example the number of exercises that have been carried out, thereby adapting a neoliberal influenced performance model. In contrast, the actions of teachers whose assessments correspond with the students' expectations are more in line with the competence model. They tend to deal with the few students who disagree by engaging in dialogue. Hyltegren (2014) focuses on the limitations of the knowledge requirements to meet teachers' needs when striving for equivalent grading, as emphasised in the performance model. The PE teachers described in article III do not negotiate the only quantitative knowledge requirement that students should be able to swim 200 metres. Instead, it is considered highly important to receive a pass grade. However, the qualitative requirements for movement are open for negotiation. Due to their weak specificity, with no clear starting point or limit being expressed, they are easier to negotiate without going against the official regulations.

Another dilemma, described in article II, is when the overall goals for values, norms and knowledge in the overall curriculum align with a competence model, and when the knowledge requirements, which prescribe knowledge as the only basis for grades, align with a performance model. When the state increased its power over the evaluative rules by introducing national knowledge requirements, adherence increased among the teachers reported on in article II. Nevertheless, a performance model influenced by a neoliberal discourse does not seem to meet the teachers' needs to motivate students to learn, adopt a physically active lifestyle and adopt the values and norms in the curriculum.

A third dilemma is that equity, as emphasised in the competence model, can be difficult to combine with neoliberal demands for equivalent grades. Students' different backgrounds imply that they bring different experiences and abilities to PE. In a competence model they can all be considered able, complement each other's abilities and have a personal value and status. In a performance model, the abilities prioritised in the knowledge requirement give students different positional status and value (Evans & Penney, 2008). The dilemma can be divided into two aspects: students balancing on the border of a pass grade and students aspiring for the highest grades. Some students are experienced by the teachers in article III as having little or no possibility to reach the knowledge requirements for movement. This could be due to a lack of support at home, such as the provision of training clothes and shoes, or the absence of movement experience, for example due to war in the home country. Teachers try to even out the playing field by obtaining training shoes from charity organisations, but are unable to compensate for society's inequalities. PE teachers handle this dilemma in different ways. Some set a fail grade, thereby adopting a performance model in which everyone is measured by the same criteria. Others set a pass grade (the lowest), or as one teacher in article I calls it, a hard-working grade, which can be interpreted as teachers adopting a competence model in order to motivate a student to keep on trying and learn according to their own abilities and circumstances. Rinne (2014) finds a tension between the official task of grading and the relational and ethical dimensions that teachers have to deal with. One ethical dilemma is when hard work does not give the required results in the knowledge requirement, in combination with the potential of the grading system to encourage a student to work harder.

The second aspect is students' aspirations for the highest grades. In article III, a sport discourse is seen to influence the teachers' perceptions of the knowledge requirements for movement qualities in the highest grades. Competitive sports correspond with the performance model in the sense that an ability hierarchy is created of positional status and value. Results are measured using clear criteria and, in the same way, guarantee comparability. In addition to the previously explored relationship between teachers and students' backgrounds in sport federations and the content and valued knowledge in PE, it is possible that the common values in sports and the performance model applied in the grading system contribute to the prevalence of the sport discourse in PE. Measurable sport results are in line with the neoliberal audit culture and are often regarded as fair by students who are familiar with the sport discourse. This raises the question of whether the differences between the health discourse and the performance model can explain why it is difficult for a health discourse to establish value in PE in a performance based grading system. It is noticeable that structuring the knowledge area of health and lifestyle in the knowledge requirements in a way that brings it more into line with a neoliberal intention to measure results makes it easier for the teacher to pay attention to when

grading (see article II). In previous research, the health discourse was found to be dominated by a physiology discourse that constructed health content as fitness training (Quennertedt, 2006). Fitness can also be measured quantitatively and is easy to adapt to a performative model.

## The advantages and disadvantages of the models

The two models have their advantages and disadvantages in terms of equivalence and equity. The dominance of one also affects the advantages of the other. When grading leans towards a performance model the equivalence of grades is likely to improve, thereby making grades a suitable selection instrument for the next education level and for the evaluation of schools. Grades are also instruments that parents can use to support their decisions in the free school choice. A precondition is that the knowledge requirements are clear enough to achieve valid grades. The possibility of a valid grading system based on national grading criteria has been questioned by Selghed (2004), Korp (2006) and Hyltegren (2014). Formulating national knowledge requirements unambiguously so that different teachers' grades can be compared is a difficult task and risks ending up as a 'check list'. To quote Sundberg and Wahlström: "Knowledge requirements focus on the visible and measurable but not necessarily on the most important aspects of learning" (2012, p. 351).

Grading in line with a performance model could affect the pedagogic freedom to individualise content and to adapt the teaching approach for students with different backgrounds. It might also narrow the concept of knowledge to the measurable and reduce teachers' professional assessments to measuring and controlling (cf. Hyltegren, 2014). The results in article III suggest that the measuring culture, with high demands for transparency, seems to promote the reproduction of a discourse of legitimate movements consisting of competitive sports and results. There is also a risk that content that is difficult to measure, such as a healthy lifestyle and dance, will be disadvantaged or limited to that which is measurable, such as fitness or being able to follow a beat.

In contrast, grading in line with a competence model could widen the concept of knowledge and enable more students to feel included in PE. The introduction of qualitative knowledge requirements, as investigated in article II, has the potential to challenge the reproduction of movement as a sport discourse, with results in competitive sports and a health discourse interpreted as fitness training. The assessment criteria in the competence model are not transparent, include many individualised assessments, focus on students' learning and depend on teachers' professional assessments. The competence model has a weaker classification of what is legitimate to assess and grade. Professional assessment, including many aspects of students' learning, has a tacit component that is difficult to articulate and can "work behind the back of students and teach-

ers, for example in the ways in which education reproduces existing social structures, divisions and inequalities" (Biesta, 2015, p. 77). It could also lead to a lessoning of validity and grades becoming poor instruments for selection to higher education and evaluation of schools. However, so far teachers' grading has not proved to be a threat to the function of grades as selection instruments for higher education. Success in higher education is predicted by teachers' grading rather than scholastic aptitude tests. Klapp Lekholm and Cliffordsson (2008) suggest that this may be due to teachers including characteristics that promote learning in the grades.

The two models co-exist in the Swedish school system in which steering documents describe pedagogy as a competence model and the grading system as a performance model, thus indicating that teachers have to struggle with the inconsistencies in the message systems of curriculum, pedagogy and assessment. According to Remesal (2011), it can cause dilemmas for teachers. She suggests that teachers' conceptions of assessment are related to the structure of the educational system and that there is a tension between the pedagogical aim to improve teaching and learning and society's need for accountability. Both aspects are present in the Swedish curriculum and teachers have to balance between them.

The combination can also be considered necessary to achieve the knowledge requirement in the performance model. The teachers in Hultqvist's (2011) study experienced a stronger influence of the performance model in the demands from school that all students should get a pass grade in combination with a stronger focus on performance. A competence model recognises the abilities that students bring to school. The said teachers therefore motivated students by developing relations and recognising every student's potential in order to help them to get a pass grade.

The teachers in the studies are not the only ones to experience tensions between different purposes. Here I would like to return to the teachers' comments on Facebook at the very beginning of the thesis. One of the teachers says that she is tearing her hair out in an attempt to help students who regard themselves as being punished by the system of aggregating different knowledge requirements into one grade. The problem could be contextualised as tensions between a competence model, with the purpose of motivating students, and a performance model, with societal needs for a well-defined aggregating system to achieve comparable grades. Another teacher asks whether it is possible to get the highest grade in the knowledge requirement of movement if your body is too heavy to perform all the movements. In the continuing debate on Facebook, the teachers come to different conclusions depending on whether they prioritise what is embedded in the competence or performance model. Those prioritising the values in the performance model argue that there cannot be different criteria for every student, while those prioritising the values in the competence model argue that every student should have the

opportunity to meet the knowledge requirements regardless of background or conditions, especially as the requirements do not mention any specific movement or activity. My conclusion is that the gap between policy and practice amongst the teachers in the secondary field is related to the tensions between the different interests and purposes influencing the steering documents.

## Further questions

Many new questions have arisen during the work with this thesis, most of which relate to the complexity of the research questions posed. It is also recognised that the thesis can only be considered as a first step and that others will need to further the work. In this section I outline some suggestions for this.

Difficulties in communicating movement qualities indicate that there is a need to develop a professional language with which to describe them, especially in a grading system that expects teachers to be transparent. The use of the RG technique facilitates the verbalisation of what is seldom expressed in words and could be a useful tool in this work. The RG technique is also possible to explore as a pedagogic tool for PE teachers to evaluate and develop their personal grading practices. Using the RG technique could facilitate a professional discussion by providing a basis for communication with students and colleagues.

Movement qualities was a new concept for the teachers participating in this study. Therefore it must be recognised that such new concept is difficult to adapt when facilities, equipment, time limitations and other restricting factors are the same. Further research is needed on how the interpretation of the concept will develop over time. Among the three knowledge areas in PE, movement has been the main focus in this thesis. The two knowledge areas of health and lifestyle and outdoor life and activities have not been attended to in any great depth. What PE teachers consider important for grades in those areas is still to be explored.

There seems to be a need amongst the teachers participating in the sub-studies to signal the value of the regulative discourse by using the message system of assessment. Can ways be found to signal the importance of and to motivate students to adhere to the regulative discourse other than using grades? How can teachers help students to be motivated internally, instead of externally by grades? Are all students aware of the values and norms in the curriculum that they are expected to acquire? One way of responding to these questions would be to find out how schools foster values and norms and how they communicate them to students without using grades as a message system.

In article II it is reported that the increased regulation created by the introduction of knowledge requirements has resulted in a stronger classification of what the teachers in the study consider legitimate to grade. This raises questions such as: Will future developments make the classification even stronger? If so, how will this affect the concept of knowledge in PE? How will it affect students' possibilities for individualised learning? What will happen with teachers' professional assessment, and how will it influence teachers' working situations? As indicated in the questions, I am concerned about what might happen if Sundberg and Wahlström (2012) are right in assuming that the use of knowledge requirements will emphasise the measurable at the expense of important aspects of learning.

The relations between the micro- and macro levels have been explored in the thesis in order to problematise how the tensions between interests at the macro level may cause dilemmas for PE teachers at the micro level. Problems have been identified that need further attention, such as the impact of the dominating discourses, favouring a competence or a performance model, teachers' assessment practices, quality in PE, equivalent grades and equity. Further conceptual tools or theoretical perspectives may offer possibilities for new analyses that may deepen the insights into teachers' grading practices in relation to the dominating discourse and the possible consequences for different groups of students. When investigating the consequences for different groups, the results presented in article IV suggest that the relations between different sociological parameters need to be addressed.

# 7. Conclusions and implications

In the thesis, several aspects have emerged that may contribute to the understanding of teachers' grading. One is that a stronger regulation of grading affects what teachers value. The stronger regulation increases the importance of the knowledge mentioned in the national requirements and can also affect the possibilities of students with different experiences and knowledge to get high PE grades. This calls for an ongoing debate about the consequences of different criteria, and what perception of knowledge should apply to the Swedish school system. Another aspect is related to the difficulties teachers have to find other interpretations than competitive sports to the new concept of movement qualities. In addition to the strong tradition of competitive sports in the school subject, a possible explanation is that the performance model advocated in the Swedish grading system stimulates the use of measurable achievements in the grading process. A third aspect is that the content and how it is presented is important for determining which groups of students that will receive high grades in PE. When the same goals are to be achieved in the same time period, the starting position becomes of greater importance.

The method and the theoretical framework have also contributed to the understanding of teachers' grading practices. The RG technique has proven useful to visualise aspects of grading that are not apparent in the official descriptions of teachers' grading assignments. Bernstein's (1996, 2003a,b) concepts have contributed to highlight dilemmas that teachers are exposed to in the actual assessment situations and that they are often left to manage on their own (cf. Prøitz , 2013). The concepts have also made it possible to relate the dilemmas teachers face in their practice to broader discourses and to identify dominant discourses not as taken for granted, but instead possible to challenge. I will elaborate on the contributions below.

Teachers' beliefs and values are often regarded as explanations for the variations in teachers' grading practices and the gap between policy and practice. By identifying dilemmas that teachers face when grading students in PE, another dimension becomes visible. Bernstein's concepts are used to identify dilemmas and to highlight the inconsistent messages sent by goals in curriculum, pedagogy and knowledge requirements about what is valued. My interpretation is that the dilemma caused by this inconsistency, together with other dilemmas that PE teachers experience when grading, reflects the

tensions between different interests at the macro level. The dilemmas can add to the understanding of the PE teachers' grading practices and put the focus on how they are impacted by social and cultural contexts. Bernstein's concepts can also help policymakers to anticipate and identify the dilemmas that emerge when developing and evaluating new grading systems.

Bernstein's concepts of the pedagogic device, combined with the competence and performance models, can contextualise teachers' dilemmas and provide a possible way of thinking, where the prevalent situation is not taken for granted but is open for discussion. Teachers can evaluate their dilemmas against the background of more overarching discourses and recognise them as part of the tensions that exist at the macro level. They can also identify the powers struggling for influence over what are to become dominant discourses.

Grading in PE has implications for equity. The competence and performance models reproduce existing inequalities in different ways. In order to come closer to achieving the compensatory intentions for education and PE, a discussion is needed about how the powers that influence the content and the ways in which the content is communicated affect different groups of students. The pedagogic device and the competence- and performance models are tools that can be used to conceptualise the powers involved.

There will always be new unpredictable situations and conflicting interests as long as society changes. Teachers cannot be left to deal with dilemmas caused by tensions between different interests at the macro level on their own. Policymakers need to recognise the dilemmas that policy generates. A professional discussion between teachers and policymakers is necessary in order to narrow the gap between policy and practice. To identify and discuss tensions between the privileged discourses could help teachers to handle new dilemmas as and when they emerge. It may also provide perspectives against what is considered thinkable and unthinkable (Bernstein, 2003b) can be evaluated. By challenging the thinkable and unthinkable it is possible for new discourses to emerge and for old ones to develop and thus contribute to the development of future grading practices in the subject of PE.

# 8. Summary in Swedish

## Bakgrund till studien

Denna avhandling handlar om betygsättning i skolämnet idrott och hälsa. Min ambition har varit att undersöka vad som värderas i lärares betygsättning. Betygsättning är ett starkt signalsystem som definierar och sänder budskap om vilken kunskap som är viktig (Bernstein, 2003a; Hay & Penney, 2013). Den svenska skolan gick 1994 från ett norm-relaterat betygsystem där eleverna jämfördes med varandra till ett målrelaterat, där de bedöms mot uppsatta kunskapsmål. Liksom i andra ämnen har betygsättning i idrott och hälsa visat bristande följsamhet till nationella betygskriterier (Tholin, 2006; Annerstedt & Larsson, 2010; Redelius et al., 2009; Skolverket, 2004). En tendens både i Sverige och i andra länder är att lärare i idrott och hälsa värderar annat än ämneskunskaper, till exempel motivation och ansträngning eller fysiska kvalitéer som styrka och uthållighet (López-Pastor et al., 2013; Redelius et al., 2009).

Om betyget i idrott och hälsa sänder ett annat budskap om vad som är viktigt att lära sig än vad kunskapskraven gör finns en risk att elevernas fokus riktas mot det som bedöms istället för mot innehållet i kunskapskraven (Redelius & Hay, 2009). Det budskap som förmedlas via betygen kan således få konsekvenser för elevernas lärande. Därför finns ett behov av att så tydligt som möjlig beskriva vad lärare värdesätter när de sätter betyg i idrott och hälsa. Att studera vad lärare värderar kan vara svårt på grund av att de ibland inte kan beskriva med ord vad det är som gör att en elev får ett visst betyg, utan hänvisar till en magkänsla (Annerstedt & Larsson, 2010; Skolverket, 2000). Magkänslan kan ses som ett uttryck för en tyst kunskap, det vill säga lärares professionella bedömning baserad på erfarenhet. Magkänslan är likväl viktig att försöka verbalisera då elever i den svenska skolan har rätt att få veta vad de förväntas lära sig för att få ett visst betyg och lärare hålls ansvariga för och förväntas kunna motivera de betyg de sätter. Tydliga lärandemål har också visat sig förbättra elevers lärande (Hattie, 2012). Att grunderna för betygen är transparenta är också viktigt för att säkerställa betygens likvärdighet

Betygsättning på samma grunder är prioriterat i den svenska skolan, men det har visat sig att det inte är självklart att eleven får samma betyg oavsett vilken skola den går i eller vilken lärare den har (Riksrevisionen, 2011). Ett problem med den förra läroplanen för grundskolan (Lpo94) var att kriterierna ansågs vaga och att lärarna fick bristfällig utbildning i hur den målrelaterade betygsättningen skulle gå till (Selghed, 2004; Tholin, 2006). En ny läroplan introducerades 2011 (Lgr11) med förväntningar på att den skulle ge förutsättningar för likvärdiga betyg genom nationella kunskapskrav, där endast kun-

skaper ligger till grund för betygsättningen. I kunskapsområdet rörelse står rörelsekvalitéer i fokus istället för prestationer i olika idrotter, vilket tidigare har visat sig influera betygsättningen (Redelius et al., 2009). Jag har därför undersökt hur implementeringen av Lgr11 har påverkat vad lärare värderar när de sätter betyg i idrott och hälsa och hur de bedömer de rörelsekvalitéer som kunskapskraven i Lgr11 föreskriver.

En ambition med det svenska skolsystemet är inte bara att betygen ska vara likvärdiga utan att också att utbildningen ska vara det. Innebörden av likvärdig utbildning är ”inte samma sak som att utbildningen skall vara lika. Tvärtom ska utbildningen anpassas till elevernas olika sociala hemförhållanden och andra förhållanden som skiftande språkkunskaper, olika kön, eventuella funktionshinder osv.” (Skolverket [SNAE], 2012, s. 12, 2014). En individanpassad pedagogik som tillåter många olika sätt att tillägna sig och uttrycka kunskap gör elevernas resultat svårare att jämföra (Gipps, 1994). Budskapen som betygen förmedlar kan också få konsekvenser för vilka som lyckas. Exempelvis kommer det innehåll som privilegieras i kunskapskraven att vara mer bekant för vissa grupper av elever. I mätningar av likvärdighet där betyg och internationella tester används som måttstock har skillnaden mellan olika grupper ökat sedan 1990-talet (Skolverket [SNAE], 2012, 2014). Ambitionen i denna avhandling är att undersöka hur olika sociologiska bakgrundsfaktorer, tillsammans och var för sig, påverkar betygen i idrott och hälsa.

Läraren har ibland utpekats som orsaken till att betygen brister i likvärdighet (Mickwitz, 2015). Sicilia-Camacho och Brown (2008, s. 103) visar på ett behov av att koppla aspekter av daglig pedagogisk praktik till bredare historiska och politiska diskurser; eller för att använda en mer sociologisk terminologi, koppla mikro till makro. Genom att analysera lärarnas betygsättning i relation till läroplan och influenser på denna vill jag vidga perspektivet från lärarna till en systemnivå.

## Syfte och frågeställningar

Det övergripande syftet i denna avhandling är att undersöka betygsättning i idrott och hälsa i grundskolans sista år. Avhandlingen baseras på fyra delstudier som presenteras i varsin artikel. Syftet i de tre första delstudierna är att undersöka olika aspekter av vad lärare värderar när de sätter betyg på sina elever i idrott och hälsa. Den fjärde delstudien fokuserar på hur olika sociologiska parametrar påverkar betygen i idrott och hälsa. Följande forskningsfrågor har ställts i delstudierna:

– Vad anser lärare i idrott och hälsa är viktigt för betyget och vilken relevans har det de säger sig värdera för de satta betygen? Delstudien presenteras i artikel I: *Exploring PE Teachers' "Gut Feelings": An Attempt to Verbalise and Discuss Teachers' Internalised Grading Criteria*.

– Hur påverkar nya nationella kunskapskrav vad lärare anser viktigt för betygen och vilka icke kunskapsrelaterade aspekter, om några, fortsätter de att värdera efter implementeringen? Delstudien presenteras i artikel II: *Swedish PE teachers struggle with assessment in a criterion-referenced grading system.*
– Vilka rörelser anser lärare i idrott och hälsa är legitima att betygsätta i förhållande till kunskapskraven och vad påverkar deras betygsättning? Delstudien presenteras i artikel III: *Swedish PE teachers' understandings of legitimate movement in a criterion-referenced grading system.*
– Hur påverkar sociologiska bakgrundsfaktorer betygen i idrott och hälsa? Delstudien presenteras i artikel IV: *Who gains? Sociological parameters for obtaining high grades in physical education.*

I den sista delstudien används nationella data från Skolverket för att analysera hur fyra bakgrundsfaktorer påverkar betygen i idrott och hälsa. Bakgrundsfaktorerna är kön, föräldrars utbildning, inflyttning till Sverige före eller efter skolstart och friskola eller kommunal skola.

Ett analytiskt syfte i avhandlingen är att analysera och problematisera lärares bedömningar i relation till läroplan och skolutveckling. Då resultaten i delstudierna diskuteras används ett läroplansteoretiskt perspektiv, och mer specifikt den brittiska utbildningssociologen Bernsteins begreppsapparat, för att synliggöra relationen mellan mikro- och makronivå.

# Teoretiskt ramverk

För att förstå hur social bakgrund påverkar skolresultaten använder Bernstein (2003a) begreppen *classification*, avgränsning på svenska, och *framing* som kan översättas till inramning. Med avgränsning menas, enligt Bernstein (2003a), hur väl separerade olika kategorier, exempelvis skolämnen är, det vill säga vad som kan räknas in i ämnet. Inramning syftar på hur kategorin, eller skolämnet, kontrolleras inom de givna ramarna. Med andra ord vem som har makten att bestämma vad, hur, när och var kunskap ska förmedlas. Med hjälp av dessa begrepp kan en pedagogisk verksamhet analyseras både med avseende på vad som kommuniceras och hur det kommuniceras. För att analysera hur starkt avgränsat det som anses legitimt att betygsätta i idrott och hälsa är använder jag begreppet avgränsning.

Bernstein beskriver läroplan, pedagogik och bedömning som tre sammankopplade *message systems* som i denna avhandling översätts till signalsystem:

> Formal educational knowledge can be considered to be realized through three message systems: curriculum, pedagogy and evaluation. Curriculum defines what counts as valid knowledge, pedagogy defines what counts as valid transmission of knowledge, and evalua-

> tion defines what counts as valid realization of this knowledge on the part of the taught. (Bernstein 2003a, s. 85)

I den första artikeln används signalsystemen för att analysera det budskap lärarna förmedlar via sin bedömning och betygsättning. För att sätta fokus på transformeringen mellan betygskriterier och betygsättning används Lindes (2012) modell med tre arenor. På den första arenan, formuleringsarenan, utformas styrdokumenten. På realiseringsarenan förverkligas innehållet i klassrummet i samspel mellan elever och lärare. Däremellan finns en transformeringsarena, där innehållet i styrdokumenten omvandlas till den pedagogiska verksamheten i klassrummet. Transformeringsarenan utforskas för att belysa vad som sker med de nationella betygskriterierna när de ska realiseras i lärarnas betygsättning.

I den andra artikeln analyseras budskapen i läroplanen och medföljande kunskapskrav och jämförs med de budskap som förmedlas i betygen. För att belysa betydelsen av normer och värderingar i betygsättningen används Bernsteins (1996) begrepp pedagogisk diskurs. Denna inbegriper både en instruktionell diskurs, som skapar specialiserade färdigheter (kunskap), och en regulativ diskurs, som skapar ordning och relationer (normer och värderingar). Bernstein (1996) hävdar att de två diskurserna inte går att skilja från varandra och egentligen utgör en diskurs:

> Often people in schools and classrooms make a distinction between what they call the transmission of skills and the transmission of values. These are kept apart as if there was a conspiracy to disguise the fact that there is only one discourse (Bernstein, 1996, s. 46).

I artikel II har jag ändå valt att analysera diskurserna separat för att synliggöra dem och göra dem möjliga att diskutera.

Det analytiskta syftet är att analysera lärares betygsättning i idrott och hälsa i relation till influenser på makronivån. Tidigare har transformeringen från läroplan till praktik diskuterats med hjälp av Lindes modell. Med hjälp av Bernsteins modell *the pedagogic device* kan analysen utökas till att även omfatta de intressen som påverkar utformandet av läroplanen och att skapa förståelse för kunskapens väg från systemnivå till elevnivå (Bernstein 2003b).

*The pedagogic device* beskriver hur kunskap produceras på produktionsfältet, också kallat *the primary field* (universitet och forskningscentra). Kunskap som produceras på produktionsfältet transformeras genom urval och prioriteringar till en pedagogisk form på rekontextualiseringsfältet (exempelvis en nationell läroplan). Olika grupper strider om inflytande över vilken kunskap som ska prioriteras och finnas med i läroplanen. Enligt Bernstein (2003b) kommer de aktörer som kämpar om inflytande över vad som ska inkluderas i läroplan och kunskapskrav från det officiella (staten) och från det peda-

gogiska (universitet, specialiserad media och andra aktörer med inflytande) rekontextualiseringsfältet.

Ytterligare en transformering sker då den priviligierade kunskapen i läroplanen transformeras till den kunskap som syns i undervisningen i interaktion med elever (Singh, 2002). Bernstein (2003b) benämner detta som en transformation från rekontextualiseringsfältet till reproduktionsfältet, också kallat *the secondary field* (verksamheten i skolan). Den här transformeringen är den som tidigare undersökts med hjälp av Lindes modell med arenor. Lindes modell överensstämmer med Bernsteins modell för hur den kunskap som värderas i skolan konstrueras.

I de transformeringar av kunskap som beskrivs i *the pedagogic device* uppstår en klyfta mellan de olika fälten, där ideologier får spelrum. Det innebär att det inte finns någon läroplan eller pedagogiskt förhållningsätt utan ideologi (Bernstein, 1996; Lilliedahl, 2015). *The pedagogic devic*e tydliggör hur olika aktörer strävar efter makt över vad som ska bedömas och betygsättas. I artikel III analyseras lärarnas uppfattningar av vad som anses legitimt att betygsätta i kunskapsområdet rörelse utifrån vilka influenser som syns i läroplanen och i lärarnas utsagor.

I kappans diskussionskapitel görs en översiktlig analys av de diskurser vars inflytande syns i lärarnas betygsättning i de tre första artiklarna. För att underlätta analysen används två modeller beskrivna av Bernstein (1996) som representerar olika värden och intentioner. I den första modellen, prestationsmodellen, har ämnet en stark avgränsning och inramning. Eleven bedöms efter tydliga kriterier (en synlig pedagogik) som beskriver vad eleverna ska uppnå för att få ett visst betyg och vad som fattas för nästa nivå. Prestationsmodellen fokuserar på skillnader, skapar hierarkier och tillskriver status och värde efter den position som uppnåtts (Evans & Penney, 2008). I den andra modellen, kompetensmodellen, har ämnet en svagare inramning och avgränsning. Eleven bedöms, men kriterierna kommuniceras inte alltid till eleven (en osynlig pedagogik). Kompetensmodellen fokuserar på likheter, igenkännande och erkännande av olika förmågor som kan komplettera varandra och alla elever tillskrivs personligt värde och status (Evans & Penney, 2008). Skolan är inte en isolerad institution, utan påverkas av utvecklingen i samhället och de värderingar som dominerar. Resultaten i avhandlingen kommer därför att diskuteras mot bakgrund av nuvarande läroplan och kunskapskrav samt de diskurser som påverkar dess innehåll, men också mot bakgrund av den lokala kontexten och lärarna som aktörer i transformeringen av kunskapskrav till betygsättningspraktik.

# Metoder

De tre första delstudierna i avhandlingen är intervjustudier som med olika fokus undersöker vad lärare i idrott och hälsa värderar när de sätter betyg. Intervjutekniken *Repertory Grid* (RG) används då den visat sig användbar för att få människor att sätta ord på sina uppfattningar, även de outsagda, om ett känt fenomen som exempelvis betygsättning (Björklund, 2008a; Fransella et al., 2004). Även om lärare inte kan formulera vad som krävs för ett högt betyg kan de ofta beskriva skillnaden mellan en elev med högt betyg och en elev med ett lägre betyg, förutsatt att det är deras egna elever som de känner väl. De intervjuade lärarna fick genom att jämföra tre av sina egna elever försöka formulera på vilket sätt två av dem var lika och hur de skiljde sig från den tredje eleven. Likheterna och skillnaderna benämns *constructs.* Constructs är lärarens egna uttryck för vad som skiljer elever med olika betyg åt och begreppet kommer att användas synonymt med lärarnas kriterier. Lärarna ombads reflektera kring sju till åtta elever i olika kombinationer. Därefter skattade lärarna varje elev på en femgradig skala efter hur väl de uppfyllde de constructs som hade genererats. Resultatet blev en matris som har analyserats med hjälp av programmet WebGrid5.

Urvalet till intervjustudierna var lärare som hade lärarexamen i idrott och hälsa, samt undervisade och betygsatte elever i grundskolans sista årskurser. Manliga och kvinnliga lärare med olika lång erfarenhet av betygsättning från skolor av olika storlek och olika upptagningsområden i två olika kommuner valdes för att få så många olika erfarenheter som möjligt representerade.

I den första delstudien intervjuades fyra lärare i idrott och hälsa med RG tekniken. Sammanlagt 32 elever diskuterades. De 86 constructs som lärarna genererade kategoriserades i fyra teman. Förutom RG intervjuerna fick de fyra lärarna i den första delstudien även skatta hur viktiga de ansåg de constructs de själva hade genererat var för betygen på en femgradig skala. *Prin Grid maps*, som bygger på en principalkomponentanalys, beräknades för att kunna analysera hur relevanta lärarens constructs var för de satta betygen. Därefter jämfördes matchningen med lärarens egen skattning.

Den andra delstudien var longitudinell i den bemärkelsen att tre av de fyra lärare som intervjuades i delstudie I, och som fortfarande arbetade i grundskolan, intervjuades på nytt fyra år efter den första intervjun. De nya intervjuerna genomfördes efter införandet av en ny läroplan (Lgr11). Även denna gång användes RG tekniken. Totalt diskuterade de tre lärarna 45 elever vid de två tillfällena. Constructen i den andra intervjun kategoriserades i samma fyra teman som i vid första tillfället. Fördelningen av de totalt 125 constructs lärarna genererat vid de två tillfällena jämfördes över tid och i relation till läroplan och kunskapskrav.

I den tredje delstudien intervjuades sju lärare i idrott och hälsa med RG tekniken. Totalt diskuterades betygsättningen av 49 elever och 150 constructs genererades. Därutöver genomfördes semistrukturerade intervjuer med samma sju lärare. I de semistrukturerade intervjuerna presenterades kunskapskraven för rörelse ett efter ett och lärarna ombads berätta hur de bedömde och betygsatte respektive kunskapskrav och vilka eventuella problem de stötte på i betygsättningen.

Den fjärde delstudien bygger på registerdata från Skolverket. Data omfattar alla elever som slutade årskurs nio 2014 (n=95317). Förutom information om avgångsbetyg i idrott och hälsa finns variablerna elevens kön, föräldrars utbildning, om eleven flyttat till Sverige efter skolstart och om eleven går i fristående eller kommunal skola. De fyra bakgrundsvariablerna är dikotoma. För att också ta hänsyn till variablernas påverkan på varandra gjordes en multivariat ordinal regressionsanalys, där effekten av en variabel kontrolleras för de andra variablerna, samt en stratifierad multivariat regressionsanalys för variabeln flyttat till Sverige efter skolstart eller bott hela skoltiden i Sverige. För att studera effekterna på betygen i idrott och hälsa för de elever som tillhör den antagna kategorin *högsta odds* för att få höga betyg med kategorin *lägsta odds* för att få höga betyg, skapades en ny dikotom variabel. De kategorier som utfallit med högsta odds för att få ett högre betyg i idrott och hälsa kombinerades (n=7350) och de kategorier som utfallit med lägsta odds kombinerades (n=2003). Därefter genomfördes en ordinal regressionsanalys på den nya variabeln.

# Sammanfattning av artiklarna

## Artikel I: Exploring PE teachers' 'gut feeling': An attempt to verbalise and discuss teachers' internalised grading criteria

Syftet var att undersöka vad fyra lärare i idrott och hälsa anser viktigt för betyget i idrott och hälsa. Studien ägde rum 2009, innan implementeringen av den nya läroplanen, Lgr11. Resultaten visar att lärarna i studien värderar motivation, kunskaper, självförtroende och sociala färdigheter när de sätter betyg i idrott och hälsa. Dessa fyra teman är alla relevanta för de satta betygen i en analys med *Prin Grid maps* trots att det i den dåvarande läroplanen, Lpo94, enbart är elevers kunskaper som ska ligga till grund för betygsättning. Lärarnas undervisningssituation kan spåras i innehållet i de tre teman som inte avser ämneskunskaper. De kriterier (constructs) som lärarna har genererat kan tolkas som att de reflekterar lärares pedagogiska och disciplinära intentioner som syftar till att eleverna ska vara aktiva, ta ansvar för sitt eget lärande och skapa en stimulerande inlärningssituation för alla. Min tolkning är att lärarna använder betygen för att förmedla budskapet att detta är viktiga beteenden att visa. En analys av temat kunskap

visar att lärarnas kriterier inte helt överensstämmer med de nationella kriterierna. En begränsande faktor kan vara strukturella ramar som tillgång till lokaler och lämpliga lärandemiljöer. De försvårar genomförande och bedömning av aktiviteter som exempelvis simning, utomhusaktiviteter och det lärarna betraktar som teoretisk kunskap då idrottshallarna ofta inte är anpassade för teorilektioner.

Med hjälp av beräknade *Prin Grid maps* kunde jag visa att lärarnas egna skattningar av sina kriteriers betydelse för betygen ibland inte stämmer överens med samma kriteriers relevans för de betyg lärarna har satt på eleverna. Det kan tyda på att lärarna inte är medvetna om alla aspekter av sin egen betygsättning. En annan förklaring kan vara att även om kriteriet inte skattas ha någon betydelse för betyget matchar det ett annat kriterium som har det. Om lärarens erfarenhet exempelvis säger att det är samma elever som har idrottserfarenhet som är duktiga i bollspel, så kan idrottserfarenhet få betydelse för det satta betyget även om det inte skattas som viktigt av läraren. Erfarenhet av matchande kriterier kan tolkas som ett uttryck för lärarna magkänsla.

Resultaten bidrar med kunskap om vilka nationella betygskriterier som lärarna har lätt att relatera till och vilka som de har svårare med. De visar också hur lärarna skapar egna kriterier som kan underlätta undervisningssituationen och hur tillgång till lärmiljöer påverkar vilka nationella betygskriterier som ges vikt i betygsättningen. Analysen illustrerar även hur enskilda kriteriers betydelse för betygen ibland förblir dolda, till och med för lärarna själva.

## Artikel II: Swedish PE teachers struggle with assessment in a criterion-referenced grading system

Syftet var att undersöka hur de nya kunskapskraven i Lgr11 påverkade vad lärarna värderar när de sätter betyg. Syftet var också att undersöka vilka kriterier (constructs) utanför kunskapstemat (om några) som lärarna fortsätter värdera. Efter implementeringen av Lgr11 genomfördes nya RG intervjuer med de tre lärare som deltog i delstudie I som fortfarande var verksamma i grundskolan.

Resultaten visar att det har skett förändringar i vad lärarna värderar. De genererar nästan dubbelt så många kunskapskriterier jämfört med i första intervjun, men endast hälften så många om motivation. Självförtroende och sociala färdigheter nämns nästan inte alls. Lärarna talar om teoretisk kunskap oftare och på ett sätt som är i linje med kunskapskraven. Det verkar som om de nya kunskapskraven hjälper lärarna att fokusera mer på kunskaper och mindre på elevernas beteende och egenskaper. Motivation kvarstår dock som underlag för betygen. Bernsteins (2003a) begrepp om de sammankopplade signalsystemen: läroplan, pedagogik och bedömning, kan erbjuda en möjlig förklaring till motståndskraften i motivationstemat. Bernstein (1996) poängterar att för-

medlingen av kunskap (den instruktionella diskursen) och värderingar (den regulativa diskursen) inte kan separeras. Läroplanen förmedlar att den instruktionella och den regulativa diskursen är lika viktiga och har mål för båda. Enligt kunskapskraven ska dock bara den instruktionella betygsättas. Betygsättning enligt kunskapskraven signalerar därmed att kunskap är viktigare än normer och värderingar, vilket är ett annat budskap än i resten av läroplanen där skolans värdegrund ges samma utrymme som kunskaper. Resultaten kan tolkas som att lärarnas tendens att betygsätta den regulativa diskursen är ett sätt att kompensera för skillnaderna i budskapen från kunskapskrav och mål i läroplanen och ge ökade möjligheter att uppnå läroplanens mål för den regulativa diskursen.

## Artikel III: Swedish PE Teachers' Understandings of Legitimate Movement in a Criterion-referenced Grading System

Syftet var att undersöka vad lärare i idrott och hälsa anser vara legitim kunskap då de betygsätter kunskapskraven för rörelse i Lgr11. Ytterligare ett syfte var att undersöka faktorer som påverkar deras bedömningar. Sju lärare i idrott och hälsa intervjuades.

Resultaten visar att lärarna ofta baserar sina bedömningar på elevers prestationer i olika idrotter och därmed också på den styrka och kondition som krävs för att prestera bra. I lärarnas utsagor syns dilemman som kan ge möjlighet att bättre förstå deras betygsättningspraktik. Ett dilemma är spänningen mellan krav på transparens om vad som betygsätts och nationella kvalitativa kunskapskrav för rörelse som kan vara svåra att förklara för elever och föräldrar. Ett annat dilemma är spänningen mellan lärarens pedagogiska ambitioner att möjliggöra för alla elever att lära sig utifrån sina egna förutsättningar och kravet på standardiserade, likvärdiga betyg. Socioekonomisk bakgrund, status i gruppen och brist på tidigare rörelseerfarenheter (det senare exemplifieras i lärares tal om nyanlända elever som kan ha begränsade rörelseerfarenheter) kan ibland göra det svårt för vissa elever att uppnå ett godkänt betyg. Det förekommer att lärare kompenserar för elevernas olika förutsättningar och anpassar kunskapskraven för att göra dem möjliga att uppnå.

Lärarna i studien bevarar ofta etablerade traditioner istället för att använda de möjligheter som de kvalitativa formuleringarna i kunskapskraven ger för att utmana dem. Lärares tolkningar av kunskapskraven styrs av ett komplext nätverk av influenser på olika nivåer och i denna studie synliggörs några av dem. Exempelvis svårigheter att kommunicera rörelsekvalitéer i kombination med krav på transparens och lärares pedagogiska intentioner att motivera alla elever till lärande i kombination med krav på likvärdiga betyg.

## Artikel IV: Who gains? Sociological parameters for obtaining high grades in physical education

Syftet var att undersöka hur olika bakgrundsfaktorer såsom kön, föräldrars utbildning, migrationsbakgrund och skolval påverkar betygen i idrott och hälsa. Skolverkets registerdata för alla elever som gick ut grundskolan 2014 (n= 95317) har använts.

Resultaten av en multivariat regressionsanalys visar att oddsen för att få ett högre betyg i idrott och hälsa är större för elever som, i fallande ordning, har bott i Sverige sedan skolstart, har minst en förälder med eftergymnasial utbildning, går i friskolor och är pojkar. Alla fyra bakgrundsfaktorerna har signifikant effekt på betygen. Stratifierade analyser av bakgrundsfaktorn migrationsbakgrund visade att skillnaden mellan pojkars och flickors odds för att få ett högre betyg i idrott och hälsa i gruppen som har flyttat till Sverige efter skolstart är större än i gruppen som har bott i Sverige under hela skolgången. Flickor med utländsk bakgrund har också visat sig mindre aktiva i idrottsrörelsen än pojkar med utländsk bakgrund (Norberg, 2016). Det tyder på att flickor med utländsk bakgrund har sämre förutsättningar i idrott både i skola och på fritiden.

För att också få en bild av hur de fyra bakgrundsfaktorerna samverkar beräknades oddskvoten för att få ett högre betyg om man kombinerar de kategorier som har högst odds och jämför med den kombination som har lägst odds med avseende på bakgrundsfaktorerna. En ny dikotom variabel skapades där pojkar som har minst en förälder med eftergymnasial utbildning, går i en friskola och har bott i Sverige sedan skolstart klassificerades som gruppen *högsta odds* (n=7350). Flickor som inte har någon förälder med eftergymnasial utbildning, går i kommunal skola och har flyttat till Sverige efter skolstart klassificerades som gruppen *lägsta odds* (n=2003). De högsta betygen, A och B, förekommer bland nästan 42 procent av eleverna i gruppen *högsta odds* vilket kan jämföra med 7 procent bland eleverna i gruppen *lägsta odds*. Gruppen *högsta odds* hade nästan sex gånger högre odds att få ett högre betyg i idrott och hälsa än gruppen lägsta odds. Resultaten visar att olika bakgrundsfaktorer interagerar med varandra.

De bestående ojämlikheter som framträder i studien tyder på att officiella strävanden att motverka betydelsen av maktstrukturer inramade av kön[6], klass och kultur inte har lyckats. Hur olika gruppers betyg gynnas eller missgynnas bär ett budskap om vilken kunskap som värderas och öppnar för ett samtal om hur ämnet idrott och hälsa kan utformas för att utmana existerande strukturer och möjliggöra att fler elever känner sig inkluderade.

[6]Bland de elever som gick ut grundskolan 2016 hade för första gången flickor något högre betyg än pojkar i idrott och hälsa. Det gäller dock inte alla grupper av flickor. Bland elever med utländsk bakgrund har pojkar fortfarande högre betyg än flickorna.

# Diskussion

Det övergripande syftet i avhandlingen har varit att undersöka betygsättning i idrott och hälsa. Jag har undersökt vad lärare värderar när de sätter betyg i idrott och hälsa samt olika elevgruppers möjligheter att lyckas i ämnet. När resultaten nu ska diskuteras inleder jag med att betrakta de likvärdighetsaspekter som behandlas i artikel IV. Därefter diskuteras resultaten i de tre första artiklarna genom att identifiera aktörer som är synliga i transformationen av läroplan till lärarnas betygsättningspraktik. Slutligen vidgas diskussionen till att även beakta den transformation som sker när läroplanen formas och de intressen och värderingar som då får betydelse. Genom att betrakta varje förflyttning av kunskap mellan olika fält som ett spelrum för ideologier (Bernstein, 1996; Lilliedahl, 2015) är intentionen att förstå lärarnas betygsättningspraktik i en vidare kontext.

## Lika villkor i idrott och hälsa

I artikel fyra studeras hur olika bakgrundsfaktorer påverkar betygen för olika elevgrupper. Resultaten visar att det är lättare för pojkar, elever i friskolor, elever med minst en förälder med eftergymnasial utbildning och elever som bott i Sverige hela sin skolgång att få höga betyg, i stigande ordning. De variabler som nämns av lärarna i intervjustudierna, socioekonomisk- och migrations bakgrund, relaterar till dem som har de högsta oddskvoterna bland de undersökta variablerna i artikel IV, migrationsbakgrund och föräldrars utbildning. Lärarna i intervjustudierna nämner också idrottsbakgrund som betydelsefullt för betygen. Tyvärr existerar inte data om elevernas idrottsbakgrund, men i en nationell utvärdering (Eriksson et al., 2005) hade självrapporterad idrottserfarenhet den högsta oddskvoten av bakgrundsfaktorerna. Vikten av idrottserfarenhet kan förstås mot bakgrund av innehållet i lektionerna och den idrottsdiskurs som speglas i lärarnas tolkning av legitim rörelse i artikel III. De som är bekanta med innehållet gynnas. Äldre pojkar, familjer med hög socioekonomisk status och barn med svensk bakgrund tenderar att vara mer aktiva i idrottsrörelsen. (Norberg, 2016; Elofsson et al., 2014). Skillnaden i deltagandet mellan pojkar med svensk och utländsk bakgrund är liten medan skillnaderna mellan flickor med svensk och utländsk bakgrund är stor (Norberg, 2016). Detta bekräftas av resultaten i artikel IV där skillnaderna i odds för att få ett högre betyg i idrott och hälsa var större mellan pojkar och flickor i gruppen som flyttat till Sverige efter skolstart, än mellan pojkar och flickor som bott i Sverige från start.

Resultaten i artikel II visar att implementeringen av kunskapskraven medfört att influenserna från en hälsodiskurs har ökat. Evans (2014) diskuterar hur idrottsämnets ansvar för hälsofrågor och barnfetma i samhället också kan innebära en exkludering och gynna de elever som uppfyller kraven på att vara hälsosamma då hälsa ges innebörden att vara smal och vältränad. Hälsodiskursen möjliggör då en reproduktion av existerande

sociala hierarkier “not despite our profession’s best intentions to address the wellbeing of children, but because of them!” (Evans, 2014, s. 321). När att vara smal och vältränad tolkas som att vara hälsosam kan grupper med hög socioekonomisk status gynnas, då de associeras med hälsosam livsstil och högre fysisk aktivitet (se Engström, 2008).

Lärares erfarenhetsbaserade förväntningar kan påverka betygsättningen (artikel I). Förväntningarna riskerar att bidra till att reproducera existerande ojämlikheter, då erfarenheterna generellt sett säger att pojkar, elever uppvuxna i Sverige, som går i friskolor och som har utbildade föräldrar lyckas bättre i idrott och hälsa (se artikel IV). Förväntningarna kan göra det svårare att upptäcka elever som inte följer mönstret. Om förväntningarna inte medvetandegörs och är möjliga att reflektera över kan de omedvetet reproducera existerande strukturer (jfr. Biesta, 2015).

Bernsteins begrepp gör det möjligt att inte bara undersöka det budskap som förmedlas utan också hur det förmedlas. Enligt Bernstein (2003b) är både synlig och osynlig pedagogik kulturellt präglade och betygen påverkas av elevers sociala bakgrund oavsett vilken pedagogik lärare använder. Han anser dock att synlig pedagogik har bättre förutsättningar att minska betydelsen av social klass för skolresultat då den troligen gör det lättare att identifiera kraven för elever som inte är vana vid kraven hemifrån (Bernstein, 2003b). För att underlätta detta föreslår han bland annat en svagare inramning av hur snabbt eleven måste lära sig saker och i vilken ordning. Kunskapskraven har istället medfört en starkare inramning av takten och ordningen på lärandet vilket kan motverka jämlikheten eftersom eleverna har olika startpositioner (Bernstein, 2003b).

Officiellt använder Sverige det Bernstein kallar ett reparationssystem för att hantera de elever som misslyckats med att uppfylla kunskapskraven. En annan strategi som Bernstein föreslår för att minska betydelsen av social klass för skolresultat är att reducera kraven på kvalitén eller kvantitet på innehållet (2003b, s. 74). Den strategin används ibland inofficiellt av lärarna i artikel III. En osynlig pedagogik är svårare för elever som inte är bekanta med det värderade sättet att tänka, tala eller röra sig. Den regulativa diskursen har inga tydliga kriterier och istället är målen formulerade i läroplanen som “Skolan ska aktivt och medvetet påverka och stimulera eleverna att omfatta vårt samhälles gemensamma värderingar och låta dem komma till uttryck i praktisk vardaglig handling” (Skolverket, 2011, s. 14). Värderingarna ska vila på kristen etik och västerländsk humanism (Skolverket, 2011, s. 9). Att Sverige har förändrats från ett relativt monokulturellt samhälle till ett multikulturellt väcker frågan om en osynlig pedagogik är tillräcklig för att förstå de ”gemensamma värderingarna” om man har andra värderingar med sig från hemmet. Då den regulativa diskursen visat sig kunna påverka betygen i delstudie I och II kan elevers olika möjligheter att förstå de värden den förmedlar påverka deras möjligheter att lyckas i idrott och hälsa.

# Aktörer i transformeringen från läroplan till praktik

I de första tre artiklarna undersöks transformeringen från läroplan och kunskapskrav till lärarens betygsättningspraktik. Olika aktörer strävar efter inflytande och de aktörer jag identifierar i tolkningarna av resultaten beskrivs här.

Lärare är uppenbara aktörer. De har möjlighet att förändra eller bevara existerande diskurser. Deras övertygelser, baserade på erfarenheter från dåtid och nutid samt förväntningar på framtiden, påverkar deras tolkningar av kunskapskraven och är synliga i de personliga kriterier de använder (artikel I och III). Lärares värderingar, övertygelser och förväntningar påverkar inte bara deras tolkningar, urval och implementering av de den förväntade bedömningen, den kan även ge ytterligare eller alternativa bidrag till fältet (Hay & Penney, 2013, s. 45).

Statens styrning av skolan kan anses öka i och med införandet av nationella kunskapskrav och därmed begränsas lärares inflytande över betygssättningen. Det kan ses som ett försök att ersätta lärares professionella bedömningar med enhetliga kriterier. Formuleringen och tydligheten i kriterierna har visat sig viktiga för om och hur väl lärarna kan följa dem (artikel II och III).

En annan grupp som har inflytande över transformeringen är eleverna. Genom att stödja eller göra motstånd mot den aktuella betygsättningspraktiken kan de påverka möjligheterna för systemet att fungera och i artikel III diskuteras hur det kan medföra anpassningar av det innehåll som ska bedömas. En majoritet av eleverna har erfarenhet av tävlingsidrott. Andra aktiviteter än traditionella föreningsidrotter kan framkalla motstånd och i artikel III ges exempel på elever som vägrar delta om de är osäkra på att de kan prestera väl. Alla elever känner sig dock inte bekväma med tävlingsidrotter och en mindre grupp har erfarenheter av marginalisering och exkludering (Eriksson et al., 2005).

Skolledning, elever och föräldrar kan ses som aktörer i ett decentraliserat skolsystem med ett fritt skolval som i Sverige. Mickwitz (2011) studie visar att lärare upplever krav från dessa tre aktörer på att ge elever höga betyg. Lärarna som deltar i delstudie III arbetar alla på kommunala skolor och nämner krav från föräldrar och elever men inte skolledning.

Olika dilemman som reflekterar intressen på makronivån blir synliga i resultaten i studierna. De tidigare beskrivna kompetens- och prestationsmodellerna kommer att användas för att beskriva olika värderingar som kan orsaka dessa dilemman. Balansen mellan modellerna kan ses som utfallet av transformeringen från produktionsfältet till rekontextualiseringsfältet.

# Prestationsmodellen och kompetensmodellen

Det analytiska syftet har varit att analysera vad lärare värderar när de sätter betyg i idrott och hälsa i relation till olika influenser som påverkar styrdokument och praktik. Biesta et al. hävdar att ”access to wider discourses about teaching and education would provide teachers with a perspective on the beliefs they and their colleagues hold, and would provide a horizon against which such beliefs can be evaluated” (2015, s. 638). På liknande sätt vill jag med hjälp av Bernsteins begrepp belysa influenser som påverkar läroplan och praktik som en horisont mot vilken lärares dilemma kan värderas. Förhoppningsvis kan det hjälpa lärare att identifiera sina problem som del av större diskurser och möjliggöra en diskussion där dominerande diskurser är möjliga att utmana och inte tas för givna.

De influenser som synliggjorts i tolkningarna av resultaten, en neoliberal diskurs, en pedagogisk diskurs, en idrottsdiskurs och en hälsodiskurs, representerar olika värderingar och syften. Här beskrivs de kort och relateras till de värden som inryms i prestations- och kompetensmodellerna (Bernstein, 1996). Beskrivningen är inte en fullständig spegling av hela det komplexa nät av influenser som påverkat utvecklingen av den svenska läroplanen utan utgör endast ett försök att kontextualisera de influenser som framträtt studien. Bilden av de två modellerna är komplex och har många variationer (Bernstein, 1996), men i analysen behålls distinktionerna och anpassas till svenska förhållanden för att göra modellerna tydliga och möjliga att diskutera.

Prestationsmodellen fokuserar på olikheter som skapar hierarkier, beroende på prestationer (Bernstein 1996; Evans & Penney, 2008). Den ses här som representerad av den neoliberala diskurs som influerat många beslut om det svenska skolsystemet de senaste två årtiondena. Diskursen förespråkar att kvalitet och effektivitet i skolan ska stärkas med hjälp av fria skolval och konkurrens mellan skolor. Avregleringen och statens minskade inflytande över hur verksamheten ska bedrivas har ersatts av en resultatstyrning av skolan där målrelaterade betyg har fått en viktig funktion för utvärdering av skolor utöver att fungera som urvalsinstrument till högre utbildning. Likvärdiga betyg har fått en central roll. Betygen på skolnivå publiceras och kan användas som marknadsföring för skolor och stöd för beslut för föräldrar och elever i det fria skolvalet. Kunskapskraven borgar för likvärdiga betyg och föreslås förmedla en kunskapssyn som säger att kunskap begränsas av kriterier som är synliga och mätbara (Sundberg & Wahlström, 2012). Lärares professionella bedömningar som ofta innehåller en tyst dimension blir ett hot mot de likvärdiga betygen. Idrottsdiskursen med tävlingsidrotter som har influerat innehållet i idrott och hälsa kan ses som en del av prestationsmodellen i den bemärkelsen att idrottsresultat ofta är mätbara, väldefinierade och jämförbara.

Kompetensmodellen förmedlar istället en syn på att alla elever är kapabla. Alla har olika kunskaper som kan komplementera varandra (Bernstein, 1996; Evans & Penney, 2008). Den anses här vara representerad av en pedagogisk diskurs som sätter elevers utveckling och lärande i centrum och som dominerar den allmänna delen av den svenska läroplanen. Undervisningen ska anpassas till elevers individuella förutsättningar och syftar till att elever förutom kunskaper utvecklar specifika normer och värderingar i enlighet med de som uttrycks i den föreskrivna värdegrunden. Bernstein menar att både den instruktionella och den regulativa diskursen är en del av den pedagogiska diskursen. I kompetensmodellen har kunskap en svag avgränsning och inramning och tolkas bredare än i prestationsmodellen. Istället för likvärdig betygsättning, som i prestationsmodellen, är fokus på en individualiserad likvärdig undervisning och elevers lärande. Lärare gör många professionella och individuella bedömningar på olika sätt, vilket gör att de inte är särskilt jämförbara (Gipps, 1994). Betygen kan fungera för att motivera elever, men kriterierna är inte transparenta. I kunskapskraven kan kompetensmodellen spåras i de kvalitativa formuleringarna för de olika betygen som ger utrymme för lärarnas professionella bedömningar. Hälsodiskursen, som har stort inflytande över kursplanen i idrott och hälsa, kan ses som en del av kompetensmodellen i den bemärkelsen att hälsobegreppet har en svag inramning som bygger på den personliga upplevelsen och ämnet idrott och hälsa syftar till elevens personliga utveckling till en fysiskt aktiv person genom hela livet.

## Kontextualisering av tre dilemman från studien

Kompetens- och prestationsmodellerna existerar parallellt i läroplanen med separata mål och värderingar som ibland sammanfaller, men som ibland också skapar dilemman för lärarna vid betygsättning. Jag kommer att ge några exempel på hur de två modellerna kan kontextualisera dilemman som framkommer i studierna. Det första dilemmat, beskrivet i artikel III, är kvalitativa kriterier som antyder en kompetensmodell i kombination med krav på transparens och ansvar för de satta betygen som kan ses som ett uttryck för en neoliberalt influerad prestationsmodell. Kvalitativt utryckta kunskapskrav vilar på lärarens professionella bedömning och exempelvis rörelsekvalitéer kan vara svåra att beskriva med ord (Nyberg & Larsson, 2014). Lärarnas bedömningar av rörelsekvalitéer kan vara lätta att ifrågasätta för föräldrar och elever som förväntar sig att läraren ska beskriva med ord exakt vad som fattas för att nå nästa betygsnivå. Lärarna i studien hanterar dilemmat olika beroende på den konkreta situationen. En lärare upplevde sig starkt ifrågasatt av elever och föräldrar som ställde krav på att han skulle försvara betygen muntligt. Han använde sig då av kvantitativa och mätbara krav och närmade sig därmed prestationsmodellen i sin betygsättning. Andra lärare vars bedömningar över-

ensstämmer med elevernas förväntningar har en position där de kan vara närmare kompetensmodellen och genom dialog hantera de få elever som de inte är överens med.

Ett annat dilemma som beskrivs i artikel II är när målen i läroplanen indikerar en kompetensmodell som värderar såväl normer och värderingar som kunskaper medan kunskapskraven förspråkar en prestationsmodell där bara specifik kunskap värderas. När staten ökade sitt inflytande över värderingsreglerna genom att introducera kunskapskrav ökade följsamheten i betygsättningen hos lärarna i studien. Den neoliberalt influerade prestationsmodellen verkar dock inte möta lärarnas behov av att motivera eleverna och att skapa en livslång, aktivt livsstil. En tolkning är att lärarna i studien betygsätter den regulativa diskursen (exempelvis motivation) för att öka möjligheten att uppnå läroplanens värdegrund som speglar en kompetensmodell.

Det tredje dilemmat, som beskrivs i delstudie III, är att den individanpassade likvärdiga utbildningen som förespråkas i kompetensmodellen och i läroplanen i vissa situationer kan vara svår att kombinera med neoliberala krav på likvärdiga betyg. I en prestationsmodell är det eleverna som har de kunskaper som prioriteras i kunskapskraven som gynnas, vilket ger eleverna olika positioner och värden. Lärarna i den här studien upplever ibland att vissa nyanlända elever och elever med låg socioekonomisk status har små, om ens några, möjligheter att nå kraven för ett godkänt betyg. Lärarna kommer fram till olika sätt att hantera dilemmat. Vissa sätter betyget F och ansluter sig därmed till prestationsmodellen där alla ska bedömas efter samma kriterier. Andra lärare sätter istället betyget E, vilket innebär att eleven når en godkänd nivå. Det kan tolkas som att de ansluter sig till en kompetensmodell där alla anses kapabla och har ett personligt värde för att motivera eleven till att fortsätta försöka och lära sig på sina egna villkor.

En idrottsdiskurs med fokus på resultat influerar lärarnas tolkning av kunskapskraven för de högsta betygen (jfr. Redelius et al. 2009). Värderingar från tävlingsidrotten passar väl in i en prestationsmodell då de bygger på en hierarki som ger status utifrån prestation. Tidigare har jag diskuterat relationen mellan lärares och elevers bakgrund i idrottsrörelsen och innehållet på lektionerna som en orsak till att en idrottsdiskurs fortfarande präglar undervisningen. Det är också möjligt att gemensamma värden i tävlingsidrott och prestationsmodellen i betygssystemet har bidragit till överlevnaden av idrottsdiskursen. Mätbara idrottsresultat är i linje med den neoliberala mätkulturen och anses ofta rättvisa av de elever som har erfarenhet från idrottsrörelsen. Det väcker frågan om skillnaderna mellan hälsodiskursen och prestationsmodellen har bidragit till svårigheterna att etablera en hälsodiskurs i ett prestationsbaserat betygssystem. Det kan noteras att när kunskapsområdet hälsa och livsstil strukturerades i kunskapskrav med en progression, som gör dem mer i linje med prestationsmodellen, blev det lättare för lärarna i artikel III att värdera och beakta dem vid betygsättningen.

Såväl kompetens- som prestationsmodellen har för- och nackdelar. När prestationsmodellens värden dominerar är det troligt att betygens likvärdighet ökar. Det finns dock en risk att det påverkar den pedagogiska friheten att individualisera innehållet i idrott och hälsa samt att anpassa undervisningen till elevernas olika bakgrund. En svag avgränsning av vad som ska bedömas äventyrar betygens likvärdighet. Å andra sidan kan en för stark avgränsning riskera att begränsa vad kunskap kan vara till det mätbara och reducera lärares professionella bedömning till att mäta och kontrollera (jfr. Hyltegren, 2014). Modellerna visar komplexiteten när styrdokumenten föreskriver en kompetensmodell för pedagogiken och en prestationsmodell för betygssättningen. Det innebär att lärarna exponeras för olika budskap från signalsystemen läroplan och bedömning.

## Slutsats och implikationer

I avhandlingen har flera aspekter framkommit som kan bidra till förståelsen av lärares betygsättning. En är att en starkare reglering av betygsättning har betydelse för vad lärarna värderar. Den starkare regleringen ökar betydelsen av de kunskaper som omnämns i kraven, men kan samtidigt missgynna elever med andra erfarenheter och kunskaper. Därför bör det föras en ständigt pågående debatt om vilka konsekvenser olika kunskapskrav kan få för olika elevgrupper och vilken kunskapssyn som ska dominera den svenska skolan. En annan aspekt är relaterad till de svårigheter lärarna har att tolka det nya begreppet rörelsekvalitéer så att tävlingsidrottens krav på kvantitativa resultat inte tillämpas i idrott och hälsa. Förutom den starka tradition av tävlingsidrotter som finns i skolämnet är en möjlig förklaring att den prestationsmodell som förspråkas i det svenska betygssystemet stimulerar användandet av mätbara prestationer vid betygsättningen. En tredje aspekt är att innehåll och hur det presenteras har betydelse för vilka grupper som får höga betyg i idrott och hälsa. När samma mål ska nås på samma tid får startpositionen en större betydelse.

Metoden och det teoretiska ramverket har också bidragit till förståelsen av lärarnas betygsättningspraktik. RG tekniken har visat sig användbar för att synliggöra aspekter av betygsättning som inte är uppenbara i de officiella beskrivningarna av lärares betygsättningsuppdrag. Bernsteins begrepp har bidragit med att synliggöra de dilemman som lärare utsätts för i konkreta bedömningssituationer och som de ofta lämnas att hantera på egen hand. Genom att tillämpa begreppen har de dilemman lärare möter i sin praktik kunnat relateras till vidare diskurser och de dominerande diskurserna identifierats som möjliga att utmana.

Nya oförutsedda dilemman kommer att uppstå i takt med att samhället förändras. En professionell diskussion mellan lärare och beslutsfattare är nödvändig för att minska

avståndet mellan styrdokument och lärares betygsättningspraktik. Att identifiera och diskutera spänningar mellan olika diskurser som kan orsaka dilemman i betygsättningen kan hjälpa lärare att handskas med nya dilemman om och när de uppstår. Kompetens och prestationsmodellen kan utgöra en fond mot vilken det Bernstein (2003b) kallar det tänkbara och det otänkbara kan bedömas. Genom att utmana det tänkbara och det otänkbara är det möjligt för nya diskurser att produceras och existerande sådana att utvecklas i skapandet av framtidens betygsättning i idrott och hälsa.

# Acknowledgements

I would like to express my deepest appreciation and gratitude to everyone who has contributed to the accomplishment of this thesis. One of the most stimulating aspects of being a doctoral student has been the meetings with you and the challenging questions discussed. More people than I can mention have supported me in different ways and I would like to thank you all.

My supervisors Karin Redelius and Jane Meckbach deserve a special mention for keeping me on the main path when I constantly found so many other interesting side-tracks. Karin, your skilled engagement and carefulness have been invaluable, and Jane, you are the most encouraging person I know. I have learned a lot. Thank you for sharing your immense knowledge and experience with joy and patience, and for your supportive friendship.

Thanks to all those who, in addition to my supervisors, have carefully read and engaged in my manuscript. Elisabeth Hultqvist, Dean Barker and Örjan Ekblom, thank you for your helpful advice at my final seminar. Claes Annerstedt and Carolina Lundqvist, thank you for your good advice during my half-time review.

The members of the Research Group of Physical Education and Sport Pedagogy (PIF-gruppen) - Håkan Larsson, Birgitta Fagrell, Jane Meckbach, Karin Redelius, Susanne Lundvall, Åsa Bäckström, Erik Backman, Bengt Larsson, Susanne Johansson, Mathias Kempe-Bergman, Britta Thedin Jacobsson, Anna Tidén, Magnus Ferry, Anna Efverström, Åsa Liljeqvist, Rickard Håkansson, Marie Graffman-Sahlberg, Bea Gibbs, Jenny Kroon, Kertin Nilsson, Eva Linghede, Gunnar Teng, Jonas Mikaels, Daniel Roe, Sabina Vesterliund, Sara Ridderlund and for the time you were with us, Lars-Magnus Engström. Thank you for all the stimulating discussions and the good times we had. A special thanks to Håkan Larsson and Susanne Lundvall for careful readings and good advice at the PIF seminar.

Thanks to all those who have contributed to the articles in various ways. To the PE teachers who participated in the studies, thank you for sharing your experiences and insightful reflections on grading. To the anonymous reviewers, thank you for your careful readings and questions that have contributed many new perspectives. To the Swedish National Agency for Education (Skolverket) and Helena Svensson, thank you for

generously providing the data for article IV and for being so helpful. Thanks too to Hans Högberg, co-writer of the fourth article, for guiding me through the world of statistics with skill and patience.

I want to thank the Swedish School of Sport and Health Sciences (GIH) and FFN for giving me the opportunity to realise this research project. Also, all the staff at who have contributed in different ways to make me feel at home at GIH - the teachers, the administration staff, the service personnel and last but certainly not least, the librarians, thank you for all your friendly help.

To my superiors and colleagues at the University of Gävle (HiG) - Nader Ahmadi, Annika Strömberg and Anne-Sofie Hiswols, thank you for engaging in my work situation and helping me to combine my two jobs. To those involved in the Higher Seminar in Educational Sciences at the University of Gävle, thank you for careful reading along with questions and advice that opened my eyes in many ways. Mikael Westling Söderström, Kajsa Jerlinder, Anna Efverström, Per-Erik Ervasti, Maria Westin, Guilherme Elcadi, Sven Blomqvist and Göran Svedsäter, without you there would be no breaks or "fika". Thank you too for reminding me that there is a life outside work. And Anna, I will be forever grateful for noticing when I needed support, for dragging me over the bumps and also for the final reading of the manuscript. Thanks too to Sue Glover Frykman, for her professional language editing services.

Life outside work has been a source of recovery, and many friends have helped me to relax from work with travels, dinners or just a chat. Thank you all. The years as a PhD student have raced by and been full of action. Without my family's support it would not have worked. Mother and Majlis, thank you for your unconditional love. David, Jacob, Elsa and my grandchild to be, you enrich my life. Thank you for discussing with me, encouraging me and making me happy. And David, thank you for reading the Swedish summary. Finally, my partner and best friend, Lasse, thank you for your patience throughout. Without your understanding, planning and advice I would not have made it. You are the best.

# References


Annerstedt, C. (2005). Physical Education and Health in Sweden. In: M. Gerber, & U. Pühse. (Eds.), *International comparison of Physical Education: Concepts, problems, prospects* (pp. 604-629). Oxford: Meyer & Meyer Sport (UK) Ltd.

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

Apple, M. W. (2005). Education, markets, and an audit culture. *Critical Quarterly*, *47*(1/2), 11-29. doi:10.1111/j.0011-1562.2005.00611.x

Bernstein, B. (1996). *Pedagogy, symbolic control and identity: theory, research, critique*. London: Taylor & Francis.

Bernstein, B. (2003a). *Class, codes and control. Vol. 3, Towards a theory of educational transmission*. London: Routledge.

Bernstein, B. (2003b). *Class, codes and control. Vol. 4, The structuring of pedagogic discourse*. London: Routledge.

Bernstein, B. & Lundgren, U.P. (1983). *Makt, kontroll och pedagogik: studier av den kulturella reproduktionen* [Power control and pedagogy: studies of the cultural reproduction]. Stockholm: LiberFörlag.

Biberman-Shalev, L., Sabbagh, C., Resh, N., & Kramarski, B. (2011). Grading styles and disciplinary expertise: The mediating role of the teacher's perception of the subject matter. *Teaching and Teacher Education, 27*(5), 831-840. doi:10.1016/j.tate.2011.01.007

Biesta, G., (2015). What is Education For? On Good Education, Teacher Judgement, and Educational Professionalism. *European Journal of Education, 50*(1), 75-87. doi:10.1111/ejed.12109

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers & Teaching, 21*(6), 624-640. doi:10.1080/13540602.2015.1044325

Björklund, L-E. (2008a). *Från novis till expert: förtrogenhetskunskap i kognitiv och didaktisk belysning* [From Novice to Expert: Intuition in a Cognitive and Educational Perspective] (Doctoral thesis). Norrköping: University of Linköping.

Björklund, L-E. (2008b). The repertory grid technique, Making Tacit Knowledge Explicit: Assessing Creative work and Problem solving skills. In H. Middleton (Ed.), *Researching*

*Technology Education: Methods and techniques* (pp. 46-69). Rotterdam: Sense Publishers.

Borell, K. & Brenner, S-O. (1997). *Att spegla verkligheten* [To reflect reality]. Lund: Studentlitteratur.

Borg, K. (2008). 'Repertory grid som forskningsmetod' [Repertory Grid as research method]. I: V. Lindberg, & K. Borg (Eds.), *Kunskapande, kommunikation och bedömning i gestaltande utbildning* (s. 137-148). Stockholm: Stockholms universitets förlag.

Brookhart, S. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice, 10*(1), 35-36. doi:10.1111/j.1745-3992.1991.tb00182.x

Brookhart, S. (2013). The use of teacher judgement for summative assessment in the USA. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 69-90. doi:10.1080/0969594X.2012.703170

Brookhart, S. (1994) Teachers' Grading: Practice and Theory. *Applied Measurement In Education, 7*(4), 279-302.

Bryman, A. (2012). *Social research methods*. Oxford: Oxford University Press.

Chan, K., Hay, P., & Tinning, R. (2011). Understanding the pedagogic discourse of assessment in Physical Education. *Asia-Pacific Journal of Health, Sport & Physical Education, 2*(1), 3-18. doi:10.1080/18377122.2011.9730340

Cross, L., & Frary, R. (1999). Hodgepodge Grading: Endorsed by Students and Teachers Alike. *Applied Measurement In Education, 12*(1), 53-73. doi:10.1207/s15324818ame1201_4

Davies, B., & Bansel, P. (2007). Neoliberalism and education. *International Journal of Qualitative Studies In Education, 20*(3), 247-259. doi:10.1080/09518390701281751

Duan, H. (2006). Privatisation, Decentralisation and Governance in Education in the Czech Republic, England, France, Germany and Sweden. In J. Zaida (Ed.), *Decentralisation and Privatisation in Education: The Role of the State* (pp. 75-97). Dordrechts: Springer.

Ekberg, J. (2009). *Mellan fysisk bildning och aktivering: en studie av ämnet idrott och hälsa i skolår 9* [Between physical education and being active. A study of the subject physical education in year 9] (Doctoral thesis). Lund : Lunds universitet, Malmö.

Elofsson, S., Blomdahl, U. Lengheden, L. &Åkesson, M. (2014). Ungas livsstil i låg- och högstatusområden, En studie i åldersgruppen 13-16 år. *Ung livsstil* (16). Stockholm: Stockholms stad: Idrottsförvaltningen. Stockholms universitet: Institutionen för socialt arbete.

Engström, L. (2008). Who is physically active? Cultural capital and sports participation from adolescence to middle age—a 38-year follow-up study. *Physical Education & Sport Pedagogy, 13*(4), 319-343. doi:10.1080/17408980802400510

Eriksson, C., Gustavsson, K., Quennerstedt, M., Rudsberg, K., Öhman, M. & Öijen, L. (2005). *Idrott och hälsa – ämnesrapport till rapport 253, Nationell utvärdering* [National Evaluation] *2003.* Stockholm: Skolverket [SNAE], Fritzes.

Evans, J. (2004). Making a difference? Education and 'ability' in physical education. *European Physical Education Review, 10*(1), 95-108. doi:10.11771356336X04042158

Evans, J. (2014). Equity and inclusion in physical education PLC. *European Physical Education Review*, 20(3), 319-334. doi: 10.1177/1356336X14524854

Evans, J., & Davies, B. (2008). The poverty of theory: class configurations in the discourse of Physical Education and Health (PEH). *Physical Education & Sport Pedagogy*, *13*(2), 199-213. doi:10.1080/17408980701852571

Evans, J., & Penney, D. (2008). Levels on the Playing Field: the Social Construction of Physical 'Ability' in the Physical Education Curriculum. *Physical Education & Sport Pedagogy, 13*(1), 31-47. doi:10.1080/17408980701345576

Fransella,F., Bell, R., & Bannister, D. (2004). *A manual for repertory grid technique*. Chichester: Wiley.

Gard, M. (2015). 'They know they're getting the best knowledge possible': locating the academic in changing knowledge economies. *Sport, Education & Society*, *20*(1), 107-121. 10.1080/13573322.2014.957177

Gipps, C. V. (1994). *Quality Assurance in Teachers' Assessment*. Paper presented at both the Annual meeting of the American Educational Research Association, New Orleans, LA, April 4-8 and The British Educational Research Association. Available from: ERIC

Guba, E.G., & Lincon, Y.S. (1994). Competing Paradigms in Quantitative Research. In N.K. Denzin & Y.S Lincon (Eds.), *Handbook of Quantitative research* (pp. 105-117). Thousand Oaks, CA: Sage.

Hansen, K. (2014). The Importance of Ethnic Cultural Competency in Physical Education. *Strategies, 27*(3), 12-16. doi.org/10.1080/08924562.2014.900462

Hassmén, N., & Hassmén, P. (2008). *Idrottsvetenskapliga forskningsmetoder*. Stockholm: SISU idrottsböcker.

Hastie, P., & Hay, P. (2011). Qualitative approaches. In K. Armour, & D. MacDonald (Eds.). *Research Methods in Physical Education and Youth Sport* (pp. 79-94). London: Routledge.

Hattie, J. (2012). *Synligt lärande för lärare* [Visible learning for teachers]. Stockholm: Natur och kultur.

Hay, P. (2012). Assessment for learning in physical education. In D. Kirk, D. Macdonald, & M. O'Sullivan (Eds), *The Handbook of Physical Education* (pp. 312-325). London: SAGE Publications.

Hay, P. J., & Macdonald, D. (2010). The gendering of abilities in Senior PE. *Physical Education & Sport Pedagogy, 15*(3), 271-285. doi.org/10.1080/17408980903150147

Hay, P., & MacDonald, D. (2008). (Mis)appropriations of criteria and standards-referenced assessment in a performance-based subject. *Assessment in Education: Principles, Policy & Practice, 15*(2), 153-168. doi:10.1080/09695940802164184

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

Hultqvist, E. (2011) Om lärarnas förändrade yrkesvillkor, *Pedagogisk Forskning i Sverige*, *16*(3), 203–213.

Hyltegren, G. (2014). *Vaghet och vanmakt: 20 år med kunskapskrav i den svenska skolan* [Vagueness and powerlessness - 20 years of national knowledge requirements in Swedish schools] (Doctoral thesis). Göteborg : Göteborgs universitet.

Kelly, G.A. (1955). *The psychology of personal constructs vol. 1. A theory of personality.* New York: W.W. Norton & Company Inc.

Kirk, D. (2006). Sport Education, Critical Pedagogy, and Learning Theory: Toward an Intrinsic Justification for Physical Education and Youth Sport. *Quest, 58*(2), 255-264.

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

Klapp Lekholm, A., & Cliffordson, C. (2008). Discrepancies between school grades and test scores at individual and school level: Effects of gender and family background. *Educational Research and Evaluation, 14*(2), 181-199. doi:10.1080/13803610801956663

Klapp Lekholm, A., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation, 15*(1), 1-23. doi:10.1080/13803610802470425

Korp, H. (2006). *Lika chanser i gymnasiet?: en studie om betyg, nationella prov och social reproduktion*.[ Same chances in upper secondary school?: a study of grades, national tests and social reproduction] (Doctoral thesis). Malmö: University of Lund, Malmö, Sweden. Retrieved from http://libris.kb.se/bib/10166306.

Kougioumtzis, K. (2006). *Lärarkulturer och professionskoder: en komparativ studie av idrottslärare i Sverige och Grekland* [Teacher cultures and professional codes. A comparative study of physical education teachers in Sweden and Greece] (Doctoral thesis). Göteborg: Göteborgs universitet.

Larsson, H. A., Fagrell, B. A., & Redelius, K. A., (2005). Kön-Idrott-Skola. *Idrottsforum.org*. Retrieved from 20051205<http://idrottsforum.org/articles/larsson/larsson_fagrell_redelius/larsson_fagrell_redelius051214.html

Larsson, H. A., Quennerstedt, M. A., & Öhman, M. A. (2014). Heterotopias in physical education: towards a queer pedagogy?. *Gender and Education, 26*(2), 135-150. doi.org/10.1080/09540253.2014.888403

Larsson, H., & Karlefors, I. (2015). Physical education cultures in Sweden: fitness, sports, dancing … learning?. *Sport, Education & Society*, *20*(5), 573-587. doi: 10.1080/13573322.2014.979143

Larsson, H., & Redelius, K. (2008). Swedish physical education research questioned—current situation and future directions. *Physical Education & Sport Pedagogy, 13*(4), 381-398. doi:10.1080/17408980802353354

Larsson, H., Fagrell, B., & Redelius, K. (2009). Queering physical education. Between benevolence towards girls and a tribute to masculinity. *Physical Education & Sport Pedagogy*, *14*(1), 1-17. doi:10.1080/17408980701345832

Lilliedahl, J. A. (2015). The recontextualisation of knowledge: towards a social realist approach to curriculum and didactics. *NordSTEP, Nordic Journal of Studies In Educational Policy*, *1*(1), 40-47.

Lincon, Y.S. & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA:Sage.

Lindblad, S., Lundahl, L., Lindgren, J., & Zackari, G. (2002). Educating for the New Sweden?. *Scandinavian Journal of Educational Research*, *46*(3), 283-303. doi:10.1080/0031383022000005689

Linde, G. (2012) *Det ska ni veta!: En introduktion till läroplans teori* [This you should know!: An introduction to the theory of curriculum]. Lund: Studentlitteratur.

Lindström, L. (2001). Från novis till mästare: En studie av bedömningskriterier i slöjd. In C. Nygren-Landgärds & J. Peltonen (Eds.) *Visioner om slöjd och slöjdpedagogik* [Visions on sloyd and sloyd education](pp. 251-280). Vasa: Åbo Akademi, Nordiskt forum för forskning och utvecklingsarbete inom utbildning i slöjd.

Lingard, B., Mills, M., & Hayes, D. (2006). Enabling and Aligning Assessment for Learning: Some Research and Policy Lessons from Queensland. *International Studies in Sociology of Education, 16*(2), 83-103.

Londos, M. (2010). *Spelet på fältet* [The play on the field] (Doctoral thesis). Malmö: Malmö högskola.

López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: a review of international literature. *Sport, Education & Society*, *18*(1), 57-76. doi:10.1080/13573322.2012.713860

Lundahl C., Hultén M., Klapp A. &, Mickwitz L. (2015). *Betygens geografi, forskning om betyg och summativa bedömningar I Sverige och internationellt*. (Delrapport från SKOL-FORSK-projektet). Stockholm: Swedish Research Council.

Lundvall, S. (2014**)**. From girl to woman: Becoming an adult; socio-cultural factors and sports participation during adolescence. In: Women and sport*: Scientific report series*. Stockholm: SISU idrottsböcker. Available at: http://www.sisuidrottsbocker.se/amnesomraden/kvinnorochidrott/

Lundvall, S. A., & Safizadeh, P. A. (2012). Att bli en del av laget. *Svensk Idrottsforskning, 21*(3), 20-24.

Lundvall, S., & Meckbach, J. (2008). Mind the Gap: Physical Education and Health and the Frame Factor Theory as a Tool for Analysing Educational Settings. *Physical Education & Sport Pedagogy, 13*(4), 345-364. doi:10.1080/17408980802353362

Lundvall, S., & Meckbach, J. (2010). For Whom and to What End?: The challenges of the subject physical education and health seen through various perspectives. *Sport Science Review, 19*(3-4), 63-76.

MacPhail, A. (2007). Teachers' views on the construction, management and delivery of an externally prescribed physical education curriculum: Higher Grade Physical Education. *Physical Education and Sports Pedagogy, 12*(1),43-60. doi:10.1080/17408980601060267

Martinsson, L., & Reimers, E. (2014). *Skola i normer* [School in norms]. Malmö: Gleerup.

McMillan, J. H. (2003). Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice. *Educational Measurement: Issues and Practice, 22*(4), 34-43. doi:10.1111/j.1745-3992.2003.tb00142.x

McMillan, J., & Nash, S. (2000). Teacher Classroom Assessment and Grading Practices Decision Making. Paper presented at the *Annual Meeting of the National Council on Measurement in Education.* New Orleans, LA, 25-27 April 2000.

Mickwitz, L. (2011). *Rätt betyg för vem? betygsättning som institutionaliserad praktik* (Licentiate thesis). Stockholm : Stockholms universitet.

Mickwitz, L. (2015). *En reformerad lärare: konstruktionen av en professionell och betygssättande lärare i skolpolitik och skolpraktik*. [School reform and the teacher. The construction of a professional and assessing teacher in policy and teacher interviews] (Doctoral thesis). Stockholm: Stockholms universitet.

Norberg, J.R. (2016). *Statens stöd till idrotten: uppföljning 2015* (CIF 2016:3). Stockholm: Centrum för idrottsforskning [Swedish Research Council for Sport Science].

Nyberg, G., & Larsson, H. (2014). Exploring 'what' to Learn in Physical Education. *Physical Education & Sport Pedagogy, 19*(2), 123-135.

OECD. (2009). *Equity and Quality in Education: Supporting disadvantaged Students and Schools*, OECD Publishing. Retrieved from *http://dx.doi.org/10.1787/9789264130852-en*

Oliver, K. L., & Kirk, D. (2016). Towards an activist approach to research and advocacy for girls and physical education. *Physical Education & Sport Pedagogy, 21*(3), 313-327. doi.org/10.1080/17408989.2014.895803

Olofsson, E. (2007). The Swedish Sports Movement and the PE Teacher 1940-2003: From supporter to challenger. *Scandinavian Journal of Educational Research, 51*(2), 163-183. doi:10.1080/00313830701191647

Penney, D. (2013). Points of tension and possibility: boundaries in and of physical education. *Sport, Education & Society, 18*(1), 6-20. doi: 10.1080/13573322.2012.713862

Penney, D., & Chandler, T. (2000). Physical Education: What Future(s)?. *Sport, Education & Society*, *5*(1), 71-87. doi:10.1080/135733200114442

Penney, D., Brooker, R., Hay, P., & Gillespie, L. (2009). Curriculum, pedagogy and assessment: three message systems of schooling and dimensions of quality physical education. *Sport, Education and Society, 14*(4), 421-442. doi: 10.1080/13573320903217125

Prøitz, T. S. (2013). Variations in grading practice - subjects matter. *Education Inquiry (Co-Action Publishing), 4*(3), 555. doi:10.3402/edui.v4i3.22629

Quennerstedt, M. (2006). *Att lära sig hälsa* [Learning Health] (Doctoral thesis). Örebro: Örebro universitet.

Redelius, K., & Hay, P. (2009). Defining, acquiring and transacting cultural capital through assessment in physical education. *European Physical Education Review, 15*(3), 275-294. doi:10.1177/1356336X09364719

Redelius, K., Fagrell, B., & Larsson, H. (2009). Symbolic capital in physical education and health: to be, to do or to know? That is the gendered question. *Sport, Education & Society, 14*(2), 245-260. doi:10.1080/13573320902809195

Redelius, K., Quennerstedt, M., & Öhman, M. (2015). Communicating aims and learning goals in physical education: part of a subject for learning?. *Sport, Education & Society*, *20*(5), 641-655. doi: 10.1080/13573322.2014.987745

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education, 27*(2),472-482. Doi: 10.1016/j.tate.2010.09.017

Rinne, I. (2014). *Pedagogisk takt i betygssamtal: en fenomenologisk hermeneutisk studie av gymnasielärares och elevers förståelse av betyg*. [The tact of teaching in grade conferences. A phenomenological hermeneutical study of upper secondary school teachers' and students' understanding of grades] (Doctoral Thesis). Göteborg: Göteborg university.

Rossi, T., & Hooper, T. (2001). Using personal construct theory and narrative methods to facilitate reflexive constructions of teaching physical education. *The Australian Educational Researcher, 28*(3), 87-116.

Samuelsson, K., & Lindblad, S. (2015). School management, cultures of teaching and student outcomes: Comparing the cases of Finland and Sweden. *Teaching and Teacher Education, 49,* 168-177. doi:10.1016/j.tate.2015.02.014

Sandahl, B. (2005). *Ett ämne för alla?: normer och praktik i grundskolans idrottsundervisning 1962-2002*. (Doctoral thesis). Stockholm : Stockholm university.

SCB (2011). SCBDOK 3.2 Befolkningens utbildning, (Befolkningsregistret, UREG) http://www.scb.se/Statistik/UF/UF0506/_dokument/UF0506_DO_2011.pdf [2016-12-12].

SCB. (2016). Register över totalbefolkningan (RTB). http://www.scb.se/sv_/Vara-tjanster/Bestalla-mikrodata/Vilka-mikrodata-finns/Registret-over-totalbefolkningen-RTB [2016-12-12].

Schenker, K. (2011). *På spaning efter idrottsdidaktik* [In search of a physical education didactics]. (Doctoral thesis). Malmö: Malmö högskola.

Seger, I. (2014). *Betygsättningsprocess i ämnet idrott och hälsa: en studie om betygsättningsdilemman på högstadiet*. (Licentiate thesis). Örebro: Örebro university.

Selghed, B. (2004) *Ännu icke godkänt: lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen* [Not yet passed: How Teachers Experience a Criterion-Referenced Grading System and what they say about its use in Swedish Secondary School] (Doctoral thesis). Lund: University of Lund.

Sicilia-Camacho, A., & Brown, D. (2008). Revisiting the paradigm shift from the versus to the non-versus notion of Mosston's Spectrum of teaching styles in physical education pedagogy: a critical pedagogical perspective. *Physical Education & Sport Pedagogy, 13*(1), 85-108. doi:10.1080/17408980701345626

Singh, P. (2002). Pedagogising Knowledge: Bernstein's Theory of the Pedagogic Device. *British Journal Of Sociology Of Education, 23*(4), 571-582.

Sivenbring, J. (2016). *I den betraktades ögon: ungdomar om bedömning i skolan*. [In the eye of the beholder, young people on a ssessment in school] (Doctoral thesis). Göteborg: Göteborg university.

Sivesind, K. (2013). Mixed images and merging semantics in European curricula. *Journal Of Curriculum Studies*, *45*(1), 52-66.

SNAE (2000). *Nationella kvalitetsgranskningar 2000* (Rapport 190). Stockholm: Author.

SNAE (2006). *Equity trends in the Swedish school system*. (Report 275). Stockholm: Author.

SNAE (2009). *Vad påverkar resultaten i svensk grundskola? En sammanfattande analys. Kunskapsöversikt.* Stockholm: Author.

SNAE (2011). *Curriculum for the Compulsory School system, the Pre-school Class and the Leisure-time Centre 2011*. Stockholm: Author.

SNAE (2012). *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid* (Rapport 374). Stockholm: Author.

SNAE (2014). *Education equity in the Swedish school system* (Report 374). Stockholm: Author.

Sundberg, D. A., & Wahlström, N. A. (2012). Standards-based curricula in a denationalised conception of education: the case of Sweden. *European Educational Research Journal, 11*(3), 342-356. doi: 10.2304/eerj.2012.11.3.342

Swedish Research Council (2016-02-10). CODEX, regler och riktlinjer för forskning. http://www.codex.vr.se/manniska6.shtml [2016-12-12].

Swedish National Audit Office (2011). *Lika betyg, lika kunskap? En uppföljning av statens styrning mot en likvärdig betygsättning i grundskolan* [Equal grades, equal knowledge?] (Rapport 2011:23). Stockholm: Riksrevisionsverket.

Swedish Schools Inspectorate. (2010). *Kvalitetsgranskning, Rätten till kunskap. En granskning av hur skolan kan lyfta alla elever* [Evaluation of quality, the right to knowledge] (Rapport 2010:14). Stockholm: Skolinspektionen.

Thedin Jakobsson, B. A., Lundvall, S. A., Redelius, K. A., & Engström, L. A. (2012). Almost All Start But Who Continue?: A Longitudinal Study of Youth Participation in Swedish Club Sports. *European Physical Education Review, 18*(1), 3-18. doi.org/10.1177/1356336X11430660

Tholin, J. (2006) *Att kunna klara sig i ökänd natur: en studie av betyg och betygskriterier - historiska betingelser och implementering av ett nytt system* [Being able to survive in an unknown environment: a study of grades and grading criteria- historical factors and the implication of a new system]. (Doctoral thesis). Borås: University College of Borås.

Tinning, R. (2000). Seeking a realistic contribution: considering physical education within HPE in New Zealand and Australia. *Journal of Physical Education New Zealand, 33*(3), 8-21.

von Greiff, C. (2009). *Lika skola med olika resurser? En ESO-rapport om likvärdighet och resursfördelning* [Equal School with Unequal Resources? An ESO Report on Equality and Resource Distribution]. Stockholm: Finansdepartementet.

Wahlström, N. (2014). Equity: Policy Rhetoric or a Matter of Meaning of Knowledge? Towards a Framework for Tracing the "Efficiency-Equity" Doctrine in Curriculum Documents. *European Educational Research Journal, 13*(6), 731-743.

Walker, B.M., & Winter, D.A. (2007). The Elaboration of Personal Construct Psychology. *The Annual Review of Psychology, 58*(1), 453-477.

Zoeckler, L. G. (2007). Moral Aspects of Grading: A Study of High School English Teachers' Perceptions. *American Secondary Education*, *35*(2), 83-102.

Öhman, M. (2007). *Kropp och makt i rörelse* [Body and power in movement]. (Doctoral thesis). Örebro: Örebro university.