



LUND UNIVERSITY

Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden

The Teachers' and Students' Perspectives

Gomez, Maria

2017

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Gomez, M. (2017). *Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden: The Teachers' and Students' Perspectives*. [Doctoral Thesis (compilation)]. Lund University.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden

The Teachers' and Students' Perspectives

MARIA DEL CARMEN GOMEZ

FACULTY OF SOCIAL SCIENCES | LUND UNIVERSITY



Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden

The Teachers' and Students' Perspectives

Maria del Carmen Gómez



LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Social Sciences, Lund University,
Sweden.

To be defended at Edens Hörsal, Lund. Date 2017-10-25 and time 10.00.

Faculty opponent

Professor Per Andersson, Department of Behavioural Sciences and Learning,
Linköping University

Organization LUND UNIVERSITY Faculty of Social Sciences Department of Sociology Division of Education Author(s) Maria del Carmen Gomez	Document name Doctoral Dissertation	
	Date of issue 2017-10-25	
	Sponsoring organization The Swedish Science Council	
Title and subtitle Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden The Teachers' and Students' Perspectives		
<p>Abstract</p> <p>This dissertation acknowledges that accountability of educators for the grades they assign is a part of the foundation for a meritocratic society, based on fairness in judging merit and, therefore, in assigning grades. In this view, any grade obtained by any student must be both reliable and valid. This, in turn, implies that the process of assessment of students' work on which the grade is based must also be both reliable and valid. A recognized method to obtain this is to base assessment of students' work on verified standards which are known and recognized by all agents with an interest in the resulting grade. Accountability for grades may then be obtained by transparency, where it is possible to show what the grades are based on, whether it is quantitative or qualitative data.</p> <p>The empirical research conducted within the scope of this dissertation used qualitative research methods, including individual face-to-face interviews with 25 active and qualified science teachers working in science programs in different schools in southern Sweden, a four week ethnographic observation in three classrooms of one school—one observation in chemistry, one in physics and one in biology, and face-to-face interviews with teachers and students who participated in the ethnographic study. The findings show mutual agreement regarding assessment and grading practices among the teachers, both from teachers' and students' viewpoints.</p> <p>From the findings in this thesis it was concluded that, formally, assessment for grading is mainly based on paper-and-pencil tests, but that there are other factors influencing the assigning of grades, such as performance expectations and pressure from different agents inside and outside the school. The students are in general neither participating in the process of deciding on what and how to assess nor any decisions about the assessment itself. The assessment is, in practice, not confined to paper-and-pencil tests, but in general not supporting the learning process during the lessons.</p> <p>Taking into account the paradigm shift from merely cognitive assessment to today's action-oriented learning and assessment, as it is established in the Swedish national curriculum, I suggest that further studies of classrooms interactions and assessment practices for grading, may benefit teachers and students, as well as the society as a whole.</p>		
Key words: Assessment and Grading practices, Learning, Natural Sciences, Upper-Secondary School		
Classification system and/or index terms (if any)		
Supplementary bibliographical information	Language English	
ISSN and key title	ISBN 978-91-7753-433-4	
Recipient's notes	Number of pages 169	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature Maria del Carmen Gomez Date 2017-09-08

Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden

The Teachers' and Students' Perspectives

Maria del Carmen Gomez



LUND
UNIVERSITY

Faculty of Social Sciences Department of Sociology
Division of Education

Coverphoto by Maria del Carmen Gomez

Painting by Eva Guillén

Copyright (Maria del Carmen Gomez)

Faculty of Social Sciences
Department of Sociology
Division of Education

ISBN 978-91-7753-433-4 (print)

ISBN 978-91-7753-434-1 (pdf)

Printed in Sweden by Media-Tryck, Lund University
Lund 2017



I had scarcely passed my twelfth birthday when I entered in the inhospitable regions of examinations, through which for the next seven years I was destined to journey. These examinations were a great trial to me. The subjects which are dearest to the examiners were almost invariably those I fancied least. I would have liked to have been examined in history, poetry and writing essays. The examiners, on the other hand, were partial to Latin and Mathematics. And their will prevailed. Moreover, the questions which they asked on both these subjects were almost invariably those to which I was unable to suggest a satisfactory answer. I should have liked to be asked to say what I knew. They always tried to ask what I did not know. When I would have willingly displayed my knowledge, they sought to expose my ignorance. This sort of treatment had only one result: I did not do well in examinations.

This was especially true of my Entrance Examination to Harrow. The Headmaster, Mr. Welldon, however, took a broad-minded view of my Latin prose: he showed discernment in judging my general ability. This was the more remarkable, because I was found unable to answer a single question in the Latin paper. I wrote my name at the top of the page. I wrote down the number of the question 'I'. After much reflection, I put a bracket round it thus '(I)'. But thereafter I could not think of anything connected with it that was either relevant or true. Incidentally there arrived from nowhere in particular a blot and several smudges. I gazed for two whole hours this sad spectacle: and then merciful ushers collected my piece of foolscap with all the others and carried, it up to the Headmaster's table. It was from these slender indications of scholarship that Mr Welldon drew the conclusion that I was worthy to pass into Harrow. It is very much to his credit. It showed that he was a man capable of looking beneath the surface of things: a man not dependent upon paper manifestations. I have always had the greatest regard for him. (p. 15).

My early life. A ROVING COMMISSION (1930). By The RT. HON. Winston S. Churchill O.M. C.H., M.P. Odhams Press Limited Long Acre, London

Dedication to Evita, Rille, Tony, Lisseth and Hany

Content

Content	9
Acknowledgement.....	11
1. Introduction	13
2. Research purpose and research questions	19
3. The current thesis	21
4. The Swedish national assessment and grading system, the national curriculum — implications for assessment.....	23
4.1. Recent changes	23
4.2. Previous grading system: norm referenced.....	23
4.3. Current grading system: criteria-referenced	25
4.4. Implications for assessment	26
4.5. Swedish national curriculum statements about the concept of knowledge and about assessment — Implications for assessment.....	27
5. Research on assessment	29
5.1. Research on teachers' current assessment practices	29
5.2. Current views of learning and assessment in science education	30
5.3. Research in the field of assessment in science education	33
5.4. The research on assessment in Sweden following changes to the national grading system —1999–2009.....	38
5.5. Criticism and limits of educational assessment	39
6. Theory	43
6.1. Different forms, purposes, and levels of educational assessment	43
6.2. The concept of assessment and grading.....	46
6.3. A sociocultural view of learning and assessment	49

6.4.	A discursive approach.....	50
7.	Method	51
7.1.	The pedagogic setting, the research methods, and the analysis of data 51	
7.2.	Analysis of the videos, sound recordings, and interviews with the teachers and students	55
7.3.	Delimitations.....	58
7.4.	Methodological considerations	59
7.5.	My experience and standpoint	60
8.	Findings.....	63
8.1.	The findings, the overall research project and foci of the three published articles.....	63
8.2.	How the findings reported in the three published articles provide answers to the research question of the thesis.....	66
8.3.	The contribution of the writers of articles I, II and III.....	68
8.4.	Summary of the articles	68
9.	Discussion	71
	References	77

Acknowledgement

I am grateful to Vetenskapsrådet—the Swedish Science Council as administered by Stockholm University—for funding for the current thesis. Thanks also to the Educational Committee of Vetenskapsrådet for giving me the opportunity to spread knowledge regarding assessment and grading at the upper secondary school level in a Swedish context. The opinions, findings, and conclusions in the thesis do not necessarily reflect the views of the supporting agencies.

I want to express my gratitude to all the teachers and students who generously contributed to the current thesis and shared their thoughts and experiences with me and my research on assessment. I devote myself throughout the thesis to your thoughts and experiences. Your openness makes it possible for me to develop the best ideas I can regarding how to improve the assessment system for maximal benefit for students and society in general.

I thank all my friends and family who supported me during the entire process of my doctoral work, specifically to my dear children and their father Ricardo. I dedicate this thesis to my daughter Evita who always believed in me, inspired me and encouraged me throughout this journey, and to Jens, my beloved husband, for his tireless and unconditional support. Your love, patience and practical help in improving the quality of my English are immeasurable.

I am very grateful to Professor Alejandro J Gallard, M., Professor and Goizueta Distinguished Chair and Director of the Georgia Center for Educational Renewal, who gave me encouragement, support, unconditional friendship and professional advice. You may not know but you are one of the most valuable people who gave me hope to finalize my work.

I am grateful to Eva Kärfve of Lund University, for her unconditional support and constructive advice. You were there for me in the academic sphere where I was the loneliest. I will never forget it!

I am grateful to Professor Anders Jakobsson, Malmö University, who gave me support, encouragement and professional advice. He supported me and cooperated in my development as a doctoral student.

I am grateful to Professor Bosse Bergstedt at Lund University for his engagement, support and contributions to make it possible to conclude my thesis.

1. Introduction

This thesis examines teachers' assessment and grading practices in science education in upper secondary schools in Sweden, and does so from the perspectives of teachers and students. The aim is to understand how teachers and students perceive teachers' assessment practices for the purposes of grading. Consistency in the assessment of students plays an important role in the school system, for its legitimacy, and for the individuals concerned. Simultaneously, assessment and grading are also issues relating to the exercise of authority, which has both a wider democratic dimension and serious consequences for the individual, as well as for society in general. The process of grading in upper secondary school has attracted some interest from researchers in recent years, and has also been a focus of public debate about education in Sweden, as there are indications that assessment by grading has flaws. Of concern are students' legal rights, grade equivalence, and issues surrounding 'grade inflation' — the difference between students' increasing grades and the simultaneous decrease in learning results as measured by both international large-scale studies and national tests. In Sweden, the assessment and grading of students is entrusted to individual teachers' professionalism and there are no instruments safeguarding a consistent standard of assessment and grading. The existing national tests were initially not aimed at the issue of assessment by grading but on learning outcomes in relation to the curriculum. Strong indications of inequality, both in educational provision and grading, as well as decreasing results in international tests of educational results have raised concern in Sweden regarding national education and grading. The National Agency for Education is reforming the national tests to better reflect the different outcomes of educational practices (Lundahl, 2010). One aim of the ongoing development process is to introduce a central calibration tool, complementary to teacher assessment and grading (Waldow, 2014). Based on several important aspects of grading, not least for pedagogical purposes, it is natural to seek knowledge of how assessment and grading is performed by teachers.

Today's meritocratic society is based on mass education where individuals have access to education and compete for positions based on, amongst other things, educational merit. From this point of view, the assignment of grades makes grading, and the assessment on which grades are based, an important factor that influences both individual opportunities and the development of society.

Educational assessment has developed in the context of social competition for professional roles, a phenomenon linked to the growing complexity of advanced industrial societies and the increasing regulation of many aspects of society. As the assessment system develops, we regulate, control and legitimate its major function: to perpetuate the social, economic and political status quo (e.g. Broadfoot, 1996). This means that the development of assessment systems and assessment practices may be understood in terms of the interaction between competence, competition and control.

The development of assessment and assessment practices provides the solution to the problem of managing the provision of mass education. Students must demonstrate competence as evidence of achievement of their goals. Students who have performed best in comparison with their fellow competitors are placed highest in the ranking, allowing them further and higher educational possibilities. This is the key to competence and competition: individuals compete on an equal basis to demonstrate their claim to competence. Grading is of major importance to educational practice and society as it plays a gate-keeping role, opening or closing doors for individuals in the labor market. This is the meritocratic basis of our modern society that allows free competition based on academic ability. This is considered the fairest starting point in the allocation of opportunities for the best and most remunerative careers.

This thesis focuses on issues related to the kind of assessment science teachers practice in their everyday classroom work, as well as in special assessment activities. The research questions are based on an interest in assessment practices as part of pedagogy, on how to learn science. Western countries have a long tradition of developing science and technology — two areas that are dependent on students being well acquainted with both the content and method of the natural sciences. Assessment also has strong pedagogic implications and clearly reflects many underlying perceptions of knowledge and skills. Thus, there are many reasons for investigating how assessment is performed.

Grading requires assessment. Assessment, however, may have aims other than grading. For example, assessment may be used in a formative way, as a

means to support learning, feeding information back to the student to highlight strengths and weaknesses in comparison with goals. The broader, and primary, meaning of assessment is to take a student aside, and simultaneously correct and guide the student's knowledge by exchanging ideas and perceptions on a subject. This meaning of assessment clearly points towards the intimate connection between assessment and learning. The part of assessment that is aimed purely towards grading is usually called summative assessment. In Sweden, where the same teacher is both teaching and grading students, assessment may simultaneously be both formative and summative.

When assessing a student, a normative element always exists, for example, knowing some facts, understanding some relationship or having the skills to perform some task. This normative element thus establishes the possible goals the student may achieve through studying. It may be wise to underscore that the normative element, the possible goal for teaching and learning, may be complex and difficult to specify, and to assess.

It is, in general, of mutual interest for both the teacher and the student, that the student achieves high grades. Therefore, educational practices tend to adapt to criteria such as achieving high grades.

The Swedish grading system was changed in conjunction with reforms that were implemented around 1994. The previous system, introduced into compulsory schooling in 1962 and the upper secondary school system a few years later, was purely normative. The previous system aimed only to stratify students in each cohort. The grades were fixed nationally within each cohort and students had to compete for grades (Skolverket, 2005). The current system is criteria referenced and each student is assigned a grade corresponding to their fulfilment of the criteria for each grade. It should be noted, however, that the grading assessment system was not influenced by a change in the principles for assigning grades. Both systems rely on assessing the fulfilment of educational goals, regardless of whether these goals were judged by norm or by criteria.

Research on assessment in Sweden by academics (e.g. Wikström, 2005; Wikström & Wikström, 2005) and the Swedish National Agency for Education (SNAE) has reported on the prevalence of 'grade inflation' in Sweden. The SNAE has described grade inflation in its report, *Likvärdig betygsättning i gymnasieskolan?* [Equivalent grading in the upper secondary school?] (Skolverket, 2009). The report shows significant discrepancies in how different teachers assess students with similar levels of knowledge. The grades a student obtains in the National Agency for Education's common national tests seem to

depend, to a large extent, on the apprehension of different teachers, as well as on which school the student attends. The conclusion drawn by the SNAE is that there are structural shortcomings regarding grade equivalence and fairness in Swedish upper secondary schools. The National Agency for Education goes as far as to say that there is no good method for measuring students' actual level of knowledge in the Swedish upper secondary school system (Swedish National Agency for Education, Report 338, 2009). The SNAE suggests that important decisions are often based on students' grades, and thus that shortcomings regarding grading equivalence render grades unreliable. This has consequences for both students and society.

In summary, there appear to be some major problems in the practice of assessment by grading in upper secondary schools in Sweden. This is problematic when grades are used as a selection instrument for employment or higher education (e.g. Wikström, 2005; Wikström & Wikström, 2005). In the future, teachers will continue to play a role in interpreting grading criteria and in performing assessment by grading. It is, therefore, crucial for the educational research community to develop an understanding of how teachers carry out assessment when grading their students.

According to the current Swedish national curriculum (introduced in 1994, modified in 2011), the concept of knowledge presupposes an active classroom dialogue about why knowledge is important for today and for the future, and how learning and the acquisition of knowledge should take place. Knowledge is a complex and multifaceted concept that can be described in different ways: as facts, understanding, familiarity, and accumulated experience. The intention of the current curriculum is that students acquire and develop not only knowledge, but also values. For example, the curriculum should promote in students a lifelong desire to learn. One implication of this is that new demands are imposed on schools' working methods and organisation. For example, the curriculum demands that traditional boundaries between different vocational subjects need to be reduced and learning should not emphasise only a single competence but also include other competences. These include ethical perspectives, to 'provide students with a foundation for and support their ability to develop personal views' (Skolverket, 2013, p. 7), and environmental perspectives, 'to provide students with insights so that they can not only contribute to preventing harmful environmental effects, but also develop a personal approach to overarching, global environmental issues' (ibid).

Moreover, the national school system is based on democratic foundations that emphasise students' agency in learning and assessment. In terms of

assessment and grading, the goals of the school are that all students, individually, ‘Take responsibility for their learning and study results, and can assess their study results and need for development in relation to the requirements of the education’ (Skolverket, 2013 p. 13).

A further value, explicitly expressed in the national curriculum and specifically highlighted in the curriculum of the upper secondary school science program, is the supposition that students should develop their ability to think critically, adopt scientific ways of thinking critically, and should work in open dialogue with both fellow students and teachers. Accordingly, the task of the school is to ensure that:

Students should develop their ability to think critically, examine facts and relationships, and appreciate the consequences of different alternatives. By these means students will come closer to scientific ways of thinking and working (Skolverket, 2013, p. 5).

Thus, the Swedish curriculum has defined knowledge as something far more than a mechanical repetition of facts, and has placed knowledge in a social context. Learning in the sciences requires students to become familiar with concepts, laws, theories, principles, conventions, and special ways of working — including understanding how scientific knowledge can be applied to social, environmental, and technological issues. One characteristic of these activities is the implication that a scientific idea should be understood dialogically, because ideas must be subject to processes of social validation (Mortimer & Scott, 2003).

This shift in conceptions of knowledge in the Swedish curriculum took place some time ago. This allowed schools time to adapt both their educational tools and assessment methodology in order to fulfil all of the different requirements on knowledge and skills, and also to assess the students accordingly. In parallel with this broadening of the concept of knowledge, the major changes to schools, which brought about decentralization and freedom of choice, have impacted on the Swedish educational system (e.g. Dahlstedt, 2007; Lidström, Holm & Lundström, 2014). This has increased teachers’ responsibility for developing — locally and in cooperation with the students — educational content, methods of learning, and also assessment principles, methods and criteria. Teachers have long been relied upon to grade their students and this reliance has thus expanded in terms of teaching methods, content, goals, and assessment of both learning and grading. Acknowledging this development makes the teacher’s role as an assessor of increasing interest.

The current research was carried out in the form of three empirical studies performed in the region of Skåne in southern Sweden. The methods, analyses, and results have previously been published in three scientific peer-reviewed articles that are included in this thesis. The research questions are based on an interest in assessment practices as part of pedagogy, on how science should be learned. Western countries have long had a tradition of developing science and technology, areas dependent on students being well acquainted with the natural sciences, both in terms of content and method. Assessment also has strong pedagogic implications and clearly reflects many underlying perceptions of knowledge and skills. Thus, there are many reasons for investigating how assessment is performed.

2. Research purpose and research questions

The focus of this thesis is to examine both teacher and student perspectives on science teachers' assessment practices for the purposes of grading in Sweden. The thesis also discusses alternative assessment approaches to support learning and student involvement in the assessment process. It examines some of the characteristics of assessment for learning, such as peer- and self-assessment. These characteristics of assessment allow students participate in the assessment process by having immediate access to the outcome of the assessment.

There is a general assumption that teachers use assessment to scaffold the learning process in everyday interactions with students and to make decisions on students during the term. Thus, modern research on assessment considers it to be an integral part of learning (e.g. Swaffield, 2008; Hayward, 2015). The process of understanding classroom-based teacher assessment requires, among other things, an understanding of how teachers and students interact and how dialogues are constructed through these interactions. Assessment may be embedded in classroom interactions between a teacher and a student and, if we are trying to understand any assessment methods, these interactions are where we must look (e.g. Broadfoot & Black, 2004; Elwood, 2006).

In order to understand teacher and student perspectives on teacher assessment the following research questions are posed:

1. What kind of assessment approach is crucial for teachers when determining a student's grade and how do teachers assess students' knowledge in terms of procedural and analytical skills, familiarity, and aptitude for critical thinking?

2. To what extent and in what ways do science teachers practice assessment during classroom interactions and what do the instructional and learning discourses look like in these situations?
3. How are students involved in the assessment process and what are teachers' and students' experiences of peer- and self- assessment?
4. How do students in the upper secondary school system experience teachers' assessment and learning as part of science education and what are students' experiences in terms of their own agency in learning and assessment during lessons?

3. The current thesis

This thesis is a compilation thesis consisting of a summary of the theoretical and empirical project work, analysis of the findings, discussion and reflections, methodological consideration and three articles published in peer-reviewed journals. Two of these articles were written in collaboration with Professor Anders Jakobsson and the author of this thesis wrote the third article alone. More detailed description about the contribution of the authors to the articles can be found in section 8.3.

The thesis project was funded by Vetenskapsrådet — the Swedish Science Council — and administered by Stockholm University and the Swedish National School for Educational Assessment. The overall research presented in the thesis engages in the following areas of importance for the Swedish National School for Pedagogic Assessment:

- Survey and audit of teachers' assessment practices.
- Peer- and self-assessment (Documentation from The Swedish National School for Pedagogic Assessment, 2007).

As with all research projects of the Swedish School for Educational Assessment, this thesis follows the fundamental principles set out by Research Ethical Principles of HSFRs [Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning].

4. The Swedish national assessment and grading system, the national curriculum — implications for assessment

4.1. Recent changes

Several major changes have been made to the Swedish educational system since the mid-1990s. These changes include a shift from state controlled education with strong governing instruments towards greater freedom for local educational work in schools. The changes include allowing schools to choose the learning content, methods for learning, assessment, and evaluation of educational work (e.g. Wikström, 2005). Changes also include a new national grading system. The governing principles have shifted from specified educational methods and content towards goal fulfilment and freedom of choice for teachers, in cooperation with students, to choose both educational methods and learning content. Since assessment and grading are closely related, and with a view to understanding the circumstances for assessment in Swedish schools today, it may be of some value here to briefly describe the changes. It should be understood that in Sweden, to a major extent, the grading of students has been (and still is) the prerogative of the teacher who performs the teaching.

4.2. Previous grading system: norm referenced

Criterion for grades works normative on the part of assessment on which grades are based. When investigating assessment practices found in today's

schools, such influences that result from any grading system necessitates some knowledge of the national grading system currently in force.

In 1994 a basic change was introduced into the national Swedish grading system, replacing norm-referenced grades with a criterion-referenced system. The previous norm-referenced system, dating back to an era of social engineering during which the first standardised tests were developed by Fritz Wigfors in 1943-1944, was introduced for compulsory schools in 1962 and for upper secondary schools a couple of years later. The norm-referenced grading system contained five grades, with the lowest being '1' and the highest '5', assigned in accordance with a standard, normal distribution curve. For each cohort, the total of all national grades was fixed to a normal distribution (Skolverket, 2005).

Since grades are the basic instrument in the selection of students most suitable for higher education, norm-referenced grading system guaranteed that the highest performing students from each cohort were selected for higher education. To allow for fair grading on an individual level, centralised nationwide tests were run each year. The results of these centralised tests were used to distribute grades between classes while still maintaining the standard distribution at the national level (Skolverket, 2005). The centralised tests were, however, not used for grading each student. This grading was left to the class teacher who was considered to have a more comprehensive understanding of a student's total accumulated skills and knowledge, as compared to one single paper-and-pencil centralised test. The teacher had the freedom to assign any grade to any student provided that the mean grade corresponded, within 0.2 grade units, to the class mean grade obtained on the centralised test. Centralised tests were held only in selected subjects.

The centralised tests were marked by the teacher, who was assisted through written marking guidelines that accompanied the test. Some cooperation between teachers was assumed and encouraged to improve consistency in judgment. The tests were compiled for each year and the precise, detailed content was classified to guarantee the reliability of the tests from one year to another. For the same reason, the test for each subject was run simultaneously all over the country to avoid any premature knowledge of the content of the test amongst teachers or students.

As can be seen, the construction of the norm-referenced system was mainly for selection within each cohort. Since the sum of each grade was fixed to the standard distribution at the national level, the grades did not give any indication of changes in learning results over time. The learning results were monitored

through other instruments, for example by analysis of results from the centralised tests run every year. It should be emphasised, however, that during this era schools were run by the state and were subject to various inspections and revisions by the National Agency for Education, which monitored educational activity at all levels. Thus, the grades did not play any role in assessing the outcomes of the educational system, which was entirely performed by other means.

4.3. Current grading system: criteria-referenced

In 1994, the norm-referenced grading system was superseded by a criteria-referenced system and the previous five grades, ranging from '1' to '5', were replaced by a system of three grades: passed 'G', passed with distinction 'VG', and passed with special distinction 'MVG'. In 2011, the number of grades was increased to five, with 'E' being the lowest, through 'D', 'C', and 'B', to 'A' the highest.

Criteria for the grades are specified in qualitative terms by the National Agency for Education. There is a notable difference in the quality of knowledge and skills between grade 'G' and the higher grades in that the higher grades always require proof of performance rather than only repetitive knowledge, such as simple facts. Knowledge is defined in terms of four dimensions: facts, proficiency, understanding, and familiarity. The reform in 2011 introduced a progression table based on the knowledge requirements between different grade levels (Skolverket, 2011).

The criteria-referenced grading system was introduced together with other major changes to the Swedish educational system during an era of political market liberalism. Governing principles such as centralism, universalism and social welfare were abandoned in favour of letting schools compete and develop in different directions, so as to give students the freedom to select their school and decide on their own education (e.g. Dahlstedt, 2007; Lidström, Holm & Lundström, 2014). In order to allow schools to develop independently, and to offer students a diverse range of types of schools that would suit their wishes, many of the basic principles of the existing system, such as a school system defined as a public good, were abandoned. Simultaneously, responsibility for the education of students from years 1 to 9, and for upper secondary school (years 10 to 12), was transferred from the state to the local ('communal') level.

The principle of the criterion-referenced grading system is that it should reflect a student's level of skill and knowledge, as established by the grade and the accompanying criteria for the grade. Thus, criterion-referenced grades may be used to evaluate the outcomes of the educational system at all levels: from individuals, classes, and schools, up to the national level.

In Sweden, criteria-referenced grades are used widely to evaluate different aspects of educational outcomes at all levels, even though the system is not well calibrated against the criteria provided for the different grades in each topic.

4.4. Implications for assessment

The most notable change to the Swedish educational system at the upper secondary school level occurred following the 1994 reforms and the introduction of a new curriculum. The 2011 curriculum represented, to some extent, a revival of the former system of differentiation, placing education programs into two main groups: one group prepared students for further studies and another for vocational activities. The grading system was adjusted to incorporate five grades from the lowest 'E' to the highest 'A', whereas the criteria were largely unchanged. The system for following up educational outcomes and calibrating grades was left intact in the form it had developed since 1994. The system incorporates national tests and test banks as methods for supporting teachers' grading, as well as a tool for comparison. Teachers are, however, still grading the students they are teaching, and the grades are still based on assessment of the student's fully accumulated skills and knowledge. Since the major school reforms during the 1960s, Sweden has relied on the individual teacher for assessment and grading. Assessment and grading have since been supported by standardising tests to calibrate the grades on a national scale. The shift from a norm-referenced to a criterion-referenced grading system, and the changes to the definition of teaching goals, do, of course, influence day-to-day work in the classroom, but the basic principle, that teachers continuously and broadly assess and grade each student based on a total of all their accumulated skills and knowledge, remains unaltered. This thesis aims to investigate how this teacher assessment, which is fundamental for grading students and for examining educational performance, is performed.

4.5. Swedish national curriculum statements about the concept of knowledge and about assessment — Implications for assessment

According to the current Swedish national curriculum, the concept of knowledge presupposes an active classroom dialogue about why knowledge is important for today and for the future, and how learning and the acquisition of knowledge should take place:

Knowledge is a complex, multi-faceted concept. Knowledge can be expressed in a variety of forms — as facts, understanding, skills, familiarity, and accumulated experience — all of which presuppose and interact with each other. Teaching should not emphasise one aspect of knowledge at the cost of another (Skolverket, 2013, p. 6).

Moreover, teaching in different subjects should ‘give students knowledge of the European Union and its importance for Sweden, as well as prepare them for a society that will have closer cross-cultural and cross-border contacts’ (ibid). For this reason, it is important to provide students with international and historical perspectives. These fundamental values and tasks for the school are essential to prepare students for a life after school and are stipulated in the curriculum as follows:

The school should make use of the knowledge and experience available in the surrounding environment, and which students have acquired from i.e. working life. The world students meet in school and the work they take part in should all help in preparing students for life after school (Skolverket, 2013, p. 6).

The natural sciences program specifically establishes this philosophy in Commentaries on diploma goals:

To be prepared for higher education studies in the natural sciences, students need to develop critical thinking and scientific approaches. (...) Students should be given the opportunity to compare the natural sciences with other sciences, and discuss differences between science and non-science (Skolverket, 2013, p. 228).

With respect to the scientific tradition, studies in science programmes require students to understand the development of ideas and theories of sciences, and to study these phenomena as elements in a historical process:

The education should contain a perspective from the history of ideas, which means that the ideas and theories of the sciences are studied as parts of a historical process (Skolverket, 2012, p. 228).

Furthermore, the curricula presuppose that our modern societies need young people to understand the role of science in questions concerning sustainable development and that the school should give students the opportunity to participate in discussions about ethical issues relating to the role of science in society:

The education should give an understanding of how science and the development of society both affect and are affected by each other and in particular highlight the role of science in questions concerning sustainable development. Students should also be given the opportunity to take part in ethical discussions of the role of science in society (Skolverket, 2012, p. 228).

Assessment of learning plays an important role in education, not only as a pedagogic tool but also by determining which features of the curriculum are to be emphasised.

5. Research on assessment

5.1. Research on teachers' current assessment practices

Previous research on assessment has shown that teacher assessment varies widely, is unpredictable and suffers from low levels of validity, in part due to the inclusion of non-cognitive criteria (e.g. Klapp Lekholm & Cliffordson, 2008, 2009). The research on current teacher assessment shows that, in general, teachers are not appropriately prepared to meet the demands of classroom assessment (e.g. Brookhart, 2004; Canal, Bonini, Micciolo & Tentori, 2012; Simms & George, 2014). Grades have become a function of several different variables, rather than a pure function of what students have learned. It would appear that grades are often measures of how well a student lives up to the teacher's expectations of a good student rather than actual measures of a student's academic achievement (e.g. Allen, 2005; Brookhart, 1993; Harris & Brown, 2009). Allen (2005) suggests that if extra-academic factors are taken into consideration in grading, then other, more appropriate, forms of assessment need to be developed. The social and cultural character of assessment has been examined by Harris and Brown (2009) who found that in order to balance the needs of society, students, and the school, teachers demonstrate divergent stakeholder interests when assessing their students. Therefore, to balance the needs of the students with the accountability of assessment and subsequent grading, teachers use different kinds of assessment for different purposes. Similar studies indicate that teachers make use of their traditions and intuitions when they assess students (e.g. Allen, 2005; Harlen, 2005; Wolf, Bixby, Glenn III & Gardner, 1991). According to these scholars, teachers are frequently conflicted about the degree to which assessment should reflect behaviour and effort as separate from achievement. In general, the research suggests that teachers work alone in their classrooms, observing, commenting on, and grading students' actions, homework, and projects, almost without professional training in assessment and without transparency towards

colleagues who might otherwise provide a useful critique (e.g. Black and Wiliam, 1998; Wolf, Bixby, Glenn III & Gardner, 1991; Wyatt-Smith, Klenowski & Gunn, 2010). Selghed (2006) has analysed teachers' views on knowledge and how teachers have understood assessment since the 1994 Swedish school reform. His research shows that teachers have disparate conceptions about assessment and that the principles on which the current assessment system is based, such as theoretical conceptions about knowledge, do not influence current grading. He found that teachers grade students as they did prior to the assessment reform of 1994 and that teachers also assess students' personalities, and that in some cases even physical attributes, such as appearance, seem to influence teachers' assessment.

5.2 Current views of learning and assessment in science education

Science education in the upper secondary school

In an ethnographic study of classroom interaction in an upper secondary school in Sweden, Beach (1999) observed different demands on students taking an economic and business program (H) and those taking a science program (NV). By comparing mathematics tasks in the science programme with program (H), he observed that the difficulty of tasks was different. According to Beach, it is not possible to avoid working in the science program of an upper secondary school if a student is to stay in the program. The science program has a mechanism that prevents non-productivity, which means that students drop out if they experience difficulties in keeping up with the workload. These findings in this study suggest that science in upper secondary schools is portrayed as demanding, intended only for high achieving students, and less available for other students.

Science education in the compulsory school

Beach's (1999) view of how science program is perceived in upper secondary schools is pertinent to the educational experiences of younger pupils at compulsory school in Sweden. Although the curriculum of the compulsory school includes natural sciences, the topic is not taught to low achieving pupils. For them, teaching is focused on the three subjects that are mandatory for

admission to upper secondary school: mathematics, Swedish, and English. This situation has been of some concern for the Swedish government and has been investigated by a government commission (The Swedish National Agency for Education, [commissioned by the Swedish government] Government decision 49, 1998-12-17). The lowering of demands on pupils, results in an increasing proportion of pupils who fail to achieve one or more of the grades in the three mandatory subjects, suggesting that the focus of the proceeding of the lowering of demands is not made for pedagogical reasons, since there is no benefit in this respect. In the new, revised curriculum for upper secondary schools (SOU, 2008, p. 27), the focus on the three mandatory subjects required for admission to upper secondary school comes at the expense of other subjects, for example history or social sciences, and is counteracted by requiring grades from eight or twelve subjects, depending on which program is chosen by the pupil. The different education program now offered are divided into two main groups, one that prepares students for further study and one that prepares them for a profession. The latter group are further divided in order to include possibilities for apprenticeships in enterprises outside of the school. The revisions may be interpreted as a formalisation of the practice already developed within upper secondary schools where different groups of pupils have access to different knowledge and standards of education, as described above in the research presented by Beach (1999). A new regulation for upper compulsory schools was announced by the Swedish Education Ministry on 31st March 2011 (SKOLFS, 2011:355; SKOLF, 2011:152). [Statutes of The National Agency of Education]. The project started in 2012, with sciences, among other subjects, should be learned in the context of national recruitment late-education in primary school higher grades. This decision may be interpreted as giving reinforcement to the idea that science teaching should separate children into different classes according to their academic ability. These educational changes suggest that when students enter upper secondary school they have already been separated into different tracks and already have their own interpretations about learning science in school.

Science education in European countries

According to the High Level Group on Science Education of the European Commission (European Commission Community Research, 2007), the decreasing interest in science studies among young people is caused by the way science is portrayed. Osborne and Dillon (2008) present the same opinion in a report to the Nuffield Foundation, and add that the problem also extends

to the assessment practices for students in science programs. According to the European Commission's Community Research report (2007), some European countries are pursuing initiatives to renew science education through inquiry-based pedagogies. However, these initiatives are limited and are not of a scale to bring about substantial impact. Science education 'often fails to provide young people with the opportunity of a cumulative development of understanding and interest' and is in danger of 'being excessively factual because of the explosion in scientific knowledge and the 'adding-on' of topics to an already excessive content base (...) it comes as no surprise that students have a perception of science education as irrelevant and difficult' (European Commission Community Research, 2007, p. 8). Recent research shows that some European countries are working hard to introduce changes in the teaching of natural sciences. But in most countries, it is still reported, for instance, that there is a gender gap in teaching and learning sciences, despite the recommendations of the report by Osborne and Dillon in 2008. The gap gender remains persistent, with males dominating in the fields of science, technology, engineering, and mathematics, and has been well documented across most OECD countries (OECD, 2009, 2012, 2013) and in Europe (European Commission, 2011). Although boys and girls are equally capable of attaining the highest scores in mathematics and science, the gap gender remains persistent at almost all levels of instruction and career stage. Recent research by the European Commission (2011) has shown the need for science and gender qualities of educational and vocational guidance to increase motivation and encourage equally the interest of boys and girls in science subjects and career accomplishment in Portugal, Spain, Norway, Finland, Poland, Great Britain, the Netherlands and Italy (European Commission, 2011). There is a consensus about the lack of basic competence and interest in mathematics and science subjects in European educational contexts (e.g. Osborne, Simon & Collins, 2003; European Commission, 2007; Osborne & Dillon, 2008). Improving science education has been a high priority in many European countries but in practice only a few countries have national or regional strategies to develop special programmes, policies or projects to improve science education instruction: 'Countries which have a general, overall strategy are Germany, Spain, France, Ireland, the Netherlands, Austria, the United Kingdom and Norway. Finland had a national strategy which ended in 2002. France is the country which has most recently "put a strategy in place"' (European Commission, 2011, p. 25). Countries with national science center and institutions promoting sciences include Spain, Portugal, France, Norway, Finland, Sweden, Great Britain and Poland. When students have already

selected science education as their priority, countries that excel in supporting gifted and talented students are Denmark, Spain and Poland.

5.2. Research in the field of assessment in science education

Traditional patterns of learning and assessment in science education

Research in teaching and learning science has identified and criticised the archetypal view of science and science education in Western societies (e.g. Aikenhead, 2006; Engström & Carlhed, 2014; Hsu & Roth, 2014; Lemke, 1990). This approach is characterised by a view of science as difficult, highly mathematical, focused on facts, linked with scientific theories and concepts, and having nothing to do with issues outside of the epistemic ground of its own discipline (e. g. Lemke, 1990; Lundqvist et. al (2012). Science teachers are themselves products of an archetypal education and they transmit these views about science from generation to generation (e.g. Aikenhead, 2006; Lemke, 1990; Nyström, 2007; Treagust & Duit, 2008). Students come to science lessons with instructional conceptions and beliefs about phenomena and concepts that have no congruence with different views of science. These conceptions and beliefs are strongly held and are often resistant to change. They are often tested through summative approaches, particularly at later stages, where there is a need to provide statewide comparisons of student achievement and where there are strong, long-standing assessment traditions. This assessment approach has been shown to be biased towards tasks relying disproportionately on a narrow range of skills in particular memorization and the ability to answer low-level questions that are situated in restricted contexts. These ways of conceiving of knowledge and assessment in science education may do students a disservice as their full range of abilities are at risk of not being recognised and rewarded. If science programs are designed only to provide content on a subject, paper-and-pencil testing might suffice for determining how much and what students know. However, promoting more engaging science, for example, problem-solving, analytical thinking, argumentation, and process skills requires developing more demanding assessment practices. Many of these complex assessment practices involve the collection of evidence of broader skills such as communication, practical tests, or observation of students' performance. These activities are difficult to perform using paper-and-pencil assessment (e.g. Mintzes, Wandersee &

Novak, 2001; Osborne & Hennessy, 2003). According to Kane, Crooks and Cohen (1999), problem-solving and process skills are better assessed by means of practical objective tests, sometimes referred to as authentic assessment, or observation of student performance, so-called performance assessment.

Performance assessment

According to Kane et al. (1999), if we want to test students' abilities to design and conduct experiments, a direct approach for assessing these abilities is to permit students to design and perform experiments. Wiggins (1989) argues that this kind of assessment involves an activity in which students are directly interested in engaging and is, therefore, said to be authentic. Some scholars find it difficult to work with the method of performance assessment, and report that performance assessment is problematical because, among other things, it requires students to perform a task that teachers must observe in order to assess them (e.g. Gott & Duggan, 2002; Roberts & Gott, 2006). Based on research in the UK, they argue that the process is time-consuming and not possible to perform within the current curriculum. They do, however, discuss alternative solutions and recognise that there is currently no means of resolving the problem of assessment of students' performance in practical work. Furthermore, Messick (1994) argues that performance assessment has a downside as the appearance of fidelity or authenticity does not necessarily imply that a proposed interpretation is valid, suggesting that this assessment method does not safely measure whether students can apply their knowledge in other contexts. In many cases, however, teachers support performance assessment because of the different qualities on learning that this kind of assessment allows both teachers and students (e.g. Harlen, 1999; Mintzes, Wandersee & Novak, 2001). Performance assessment is also endorsed by many other researchers (e.g. Harlen, 1999; Osborne & Hennessy, 2003) who claim that the development of student skills is best pursued by giving them opportunities to conduct investigative practical work that will help improve their ability to interpret and present evidence, and providing them with an awareness of scientific approaches to problem-solving. Accordingly, these outcomes should be assessed as much as any factual knowledge of science. However, they stress a problem associated with the science curriculum: the requirement to simultaneously undertake practical work and to assess students' competence and skill in this domain.

The ‘Initiation–Reply–Evaluation’ (I–R–E) pattern in the classroom

A traditional pattern of science learning and assessment is the approach referred to as the I-R-E pattern, which seems to be one of the more usual kinds of classroom interactions for evaluating students’ understanding of the subject (e.g. Cazden, 2001; Lemke, 1990; Mehan, 1979; Mortimer & Scott, 2003; Wells & Mejía Arauz, 2006). The pattern has been described as two rounds of ‘Initiation–Reply–Evaluation’ (I–R–E). In such a sequential pattern of interaction, the teacher initiates an episode by asking a question with a known answer, students reply with bids for correct responses, and the teacher evaluates the responses and may initiate another round. According to these scholars, classroom interactions, such as I–R–E, are part of a complex cultural school system that, when functioning smoothly, allows participants to coordinate actions whilst achieving multiple goals. They argue that, among other things, it allows teachers to simultaneously maintain a high degree of control in the classroom whilst developing conceptual understanding and orchestrating a description of those concepts using students’, as well as the discipline’s, discourses, thus bringing them closer to grasping a set of specified concepts (Lemke, 1990). The I–R–E pattern of learning is well understood, although research suggests some shortcomings of this approach as it may constrain student participation in science discourses as well as the spontaneous inquiries of students. When using the I–R–E approach, the social atmosphere in the classroom is, according to Cazden (2001), adapted for the traditional transmission-oriented teaching of the curriculum involving theoretical knowledge. Such patterns of classroom interactions are cultural devices that help students and teachers to accomplish everyday activities, but they are not well adapted to scientific inquiry (Rop, 2002, 2003). The I–R–E pattern is neither well adapted to some pedagogic method nor is it controlled by students (Ford & Wargo, 2012; Herrenkohl & Guerra, 1998). It has been found that students’ participation in classroom discussions greatly increases when teachers do not assess students’ performance (e.g. Ollin, 2008; Radford, Ireson & Mahon, 2006). Students know that the teachers will assess them when they are talking, and may fear a poor assessment. This suggests a bias toward talking in the classroom when student participation in a class is a criterion for assessment of students’ performance. In this situation, teachers may equate listening with non-participation and it is thus suggested that the purpose of silence should be interpreted as more than an absence of talk, because it can include broader aspects of human communication, for example, visual and spatial as well as vocal (Ollin, 2008).

Learning and assessment that pays attention to the different views of science

The contemporary picture of what counts as science learning, and how to assess students' outcomes in science, may be viewed in a range of different ways (e.g. Duit, 2007; European Commission Community Research, 2007; Hsu & Roth, 2014; Osborne & Dillon, 2008; Tytler, 2007). Views and ideas about what school science should look like are diverse. Some of these ideas take account of issues such as life-long learning, environmental, and humankind issues, whereas others argue for the promotion of forms of scientific literacy (Sadler, 2009) and the nature of science (NOS) (Hodson, 2009) or features of science (FOS) (Matthews, 2012) in the curricula. There is a growing worldwide emphasis on the development of so-called key skills or life skills, regarded as critical to processes of lifelong learning which stress that the measuring of science process skills is a challenge to current directions in state and national assessment practice (e.g. Harlen, 1999). Harlen argues that teachers should include the development and achievement of these important outcomes when assessing learning in science. Other scholars claim that the development of science and technology, leading to changes in the relationship between science and society, calls for changes in the way that science is portrayed in school (e.g. Corrigan, Dillon & Gunstone, 2007; Duit, 2007; Osborne & Dillon, 2008; Treagust & Duit, 2008). According to them, these changes involve different views of science education, that is, giving students a more holistic picture of science whilst focusing less on the detail and more on broader explanatory issues — involving, among other things, the recognition of the importance of informal learning and broadening the purposes of school science (see also Osborne & Hennessy, 2003; Pedretti, 2002; Tytler, 2007).

Assessment in science education by understanding students' discourses — sociocultural approaches

Research on language as an important tool for assessing students' understanding has suggested approaches to assessing students in science education (e.g. Anderson, Zuiker, Taasoobshizazi & Hickey, 2007; Gee, 2003; Leung & Mohan, 2004; Sampson & Clark, 2008). Other approaches include questions about how students can best acquire new knowledge through shared meaning (e.g. Carter, 2007; Hargreaves, 2007; Hsu & Roth, 2014; Lemke, 1990; Yin et al., 2008; Wyatt-Smith et al., 2010). Some of these approaches are embedded in longer-term activity and may involve teacher judgment and

moderation, and student participation in the process of learning and assessment. Other linguistic approaches involve dynamic, formative, or multimodal assessment (e.g. Anderson et al., 2007; O'Byrne, 2009; Poehner, 2011; Poehner & Lantolf, 2005) and specifically propose ways to manage interactions in the classroom in order to assess students' skills by understanding students' discourses (e.g. Jakobsson et al., 2009; Johnson, 2003; Leung & Mohan, 2004). The research examines sociocultural influences on learning from the perspective of 'knowing' as a form of social activity. The use of language to assess students' understanding is by no means easy or problem-free. Many scholars report problems when teachers use language approaches to assess students' understanding (e.g. Ash et al., 2007; Lemke, 1990; Mortimer & Scott, 2003). This is partly due to the diverse issues pertaining to the interactions process, understanding the language people speak, and the epistemic tensions that arise between every day and scientific academic discourse; and partly because there are inherent tensions between the authoritative and dialogic approaches that characterise discursive interactions in the classroom (e.g. Scott, Mortimer & Aguiar, 2006).

Both the research on assessment (Black & Wiliam, 2006) and the research that focuses specifically on assessment in science education suggest that the lesson, like all social activity, is contingent and explicit only in retrospect (Lemke, 1990). This follows Wittgenstein's (1953) approach to language, which regards communication as a kind of game. The contingent nature of discourse in the classroom does not mean that discourses are built by randomly choosing words and statements. On the contrary, many rules and tools, as in games, guide the formulation of arguments and hypotheses as well as teacher directions to students, student responses, and so forth. According to Lemke (1990), in order to evaluate students' achievement in an activity, students should learn the discourses of the subjects (see also Ash et al., 2007; Gee, 2003). Similar research on discourse as a game (e.g. Anderson et al., 2007; Wells & Arauz, 2006) shows that participation in social forms of scientific discourses supports both learning and consequent performance in science classes. As is the case in games, students may be able to participate in certain types of communication activities but unable to take part in others. According to Wells and Arauz (2006), however, this requires curricula adapted to an inquiry orientation. The focus in this assessment approach is not direct assessment of the dialogue per se but rather the activities that would be likely to generate dialogue.

5.3. The research on assessment in Sweden following changes to the national grading system —1999–2009

A mapping of the research on assessment in Sweden provides an overview of assessment in primary and secondary schools across the period 1999–2009 (Forsberg & Lindberg, 2010). The mapping includes doctoral theses and Swedish, Nordic, as well as international research, on Swedish conditions. According to this survey, the research on assessment focuses on compulsory schooling. Most of the studies related to assessment research are about student performance and schools' performance, the latter being the largest category. The mapping does not report any studies of what methods the teachers use when they assess students for grading, and it does not include any study on classroom practices regarding assessment by grading or classroom practices in upper secondary schools in the field of science. Thus, the current thesis points towards a lack of studies about teachers' assessment practices in upper secondary schools, particularly in the domain of science education. The research over this period suggests that, even though changes have occurred in the Swedish national educational system and in international research on assessment, assessment approaches have not changed:

Oral or written homework, tests and examinations are common forms of assessment. They also include assessments linked to the phenomena of hands up, group steering and pilotage. The same applies to the teacher's continuous communication with the students in accordance with the united well-established pattern of question-answer- response (Forsberg & Lindberg, 2010, p. 37).

Thus, in this survey, assessment methodologies such as I–R–E discourse patterns and written examinations appear to be the most common assessment approaches used by teachers. This is a well-known but not necessarily complete description of teacher's assessment by grading. These well-known methods may well be accompanied by less formal methods of assessment. To understand teachers' assessment by grading it appears to be necessary to also search for other possible methods that teachers use. Furthermore, it is not known how curriculum requirements on the use of assessment as a tool for learning, or as a tool for encouraging student responsibility and participation, are accounted for by current assessment practices. For example, how should critical thinking, argumentation skills or the knowledge acquired outside of the classroom be assessed? This thesis aims to shed some light on this area of

research by examining science teachers' assessment practices and teachers' and students' perceptions of teachers' assessment by grading.

5.4. Criticism and limits of educational assessment

Criticism of current forms of educational assessment

There is widespread criticism directed at the traditional psychometric approach to assessment. This criticism is based on a number of factors: assessment emphasises comparisons between students rather than the specific and changing levels of attainment by individuals; there is frequently a mismatch between curricula and test content; there is pressure to test a relatively limited number of aspects of the program of instruction; and there is an assumption that students learn in a linear fashion and therefore must be taught and assessed on the basics before proceeding to more complex intellectual tasks (Broadfoot, 1996; Gipps, 1999). Many researchers in the field of assessment argue that a traditional psychometric approach to assessment has acceptable psychometric properties and is easy to score, but this approach has been criticised because it does not reflect what students really are capable of and the quality of knowledge, for example, the degree of understanding of a topic. Other reasons for the increased disapproval of traditional assessment based only on psychometric measurement, as for example objective testing, is the test's undesirable 'wash-back' effect on instruction (e.g. Broekkamp & Van Hout-Wolters, 2007; Crooks, 1988). Moreover, there is a tendency for teacher-made tests to place emphasis on lower cognitive levels (Crooks, 1988), and the undesirable impact on learning (e.g. Broekkamp & Van Hout-Wolters, 2007; Ecclestone & Pryor, 2003). The research has also found issues of ethnic bias in assessment of students (e.g. Gipps, 1999; Huston, 2005; Smith & Hawkins, 2011). Current assessment is also problematic because it defines students' attitudes in relation to their school work. For example, it influences students' sense of ownership, control of learning, and process of identity construction (e.g. Ecclestone & Pryor, 2003; Rasmussen & Friche, 2011).

The limits of educational assessment

Assessment has the tendency to measure factual and procedural knowledge that is relatively constrained in form and content and that fails to effectively identify complex knowledge. Current assessment approaches are linked to earlier theories that assumed individuals had rigid dispositions to perform work in certain ways and situations, and the characterisation of learning as a step-by-step accumulation of facts, knowledge, and skills (e.g. Gardner, 2006; Huot, O'Neill & Moore, 2010; Pellegrino, Chudowsky & Glaser, 2001). These assessment approaches lead us to ask questions about the authenticity of the inferences one can draw from students' outcomes and the need to reflect on the limits of such assessment practices. The static nature of current assessment is also criticised for limiting the development of students' learning (e.g. Pellegrino et al., 2001). It is claimed that static assessment provides snapshots of achievement at a particular time and does not pay attention to the development of students' conceptual understanding, which is at the heart of learning. Most current modes of assessment do not correspond with contemporary underlying theoretical views about how learning works and are not designed to capture the complete progression of students' learning.

Other limitations of current assessment practice are related to the issue of what new knowledge is. Learning is not only the acquisition of new knowledge but is also the interaction between new knowledge and students' prior knowledge (e.g. Davis, 1998; Yin et al., 2008). Davis (1998) argues that it is not easy, even for proficient teachers, to match the differing needs and abilities of individuals with appropriate objectives and methods. He asserts that when attempting to understand how to match one need to a pupils' understanding and skills so that it may be effectively assessed, it is not possible to determine what is new knowledge, unless we are familiar with what was the prior knowledge. One implication of this reasoning is that the acquisition of new knowledge requires prior knowledge to 'be detected with at least some semblance of accuracy' (Ibid, p. 122). This is, however, not always done and not easily accomplished. This reasoning is in line with that of other assessment specialists (e.g. Glaser & Silver, 1994; Newton, 2007). Furthermore, current assessment does not pay attention to the social character of the acquisition of knowledge. By drawing on Vygotsky, Daniels (2008) maintains a view that knowledge and learning is not located in the head of the individual but outside of it, between individuals (see also Elwood, 2006). Thus, assessment cannot identify students' abilities from what is in their heads. According to Elwood (2006), if we look within the student for their learning and knowledge we are

looking in the wrong place. From a sociocultural approach, assessment describes the relationship between the learner, the teacher, and the assessment task in the social, historical, and cultural context in which it is carried out (e.g. Greeno, More & Smith, 1993; Lave & Wenger, 1991). Greeno, More and Smith (1993) argue that it is not enough to consider that knowledge is influenced by context, they also argue that knowing, reasoning, and understanding are the relations between the cognitive agent and their situations and, therefore, it is meaningless to try to determine what individuals know separately from the context.

6. Theory

6.1. Different forms, purposes, and levels of educational assessment

The research in this field describes three levels of assessment purpose (e.g. Newton, 2007). Those levels of assessment are the judgment level, the decision level, and the impact level. Newton explains that an assessment system that is fit for one purpose will not necessarily be fit for all purposes. Assessment at the judgment level involves the technical aim of an assessment event, for example, the purpose of which is to obtain a standard-referenced judgment, expressed as a grade. It represents an assessed person's knowledge, skill or understanding, that is, an account of their competence at the end of the term. Assessment at the decision level refers to the use of a judgement assessment. Action at the decision level enables the assessment to support, for example, a selection for entry to higher education. This is the most significant usage of the term 'assessment purpose' because this level seems to be the one that is most frequently associated with it in research (Newton, 2007, p. 150). Assessment at the impact level is the assessment used to support learning. It involves: (a) the planned impacts of running an assessment system that ensures that students continue to be motivated and (b) it ensures that all students acquire a common core for each subject. To ensure that students remain motivated, assessment at the impact level requires that all students learn a common core of subjects, the assessment may well be administered on a unity-by-unity basis and aligned to a national curriculum, but this is not a prerequisite for assessment at the impact level.

Features of assessment at the judgment level

In a general sense, judgement refers to the overall outcome from diverse assessment events in the classroom. It represents the assessed student's knowledge, skill, or understanding and represents the student's competence. The judgement must be comprehensive and based on discussions with peers and an evaluation by the teacher based on a portfolio with a descriptive judgement of the student's skills. The judgement might be expressed in the form of a grade and based on evidence of performance across a long continuum.

Features of assessment at the decision level

The use of knowledge measurement at the decision level has followed two historical lines of occurrence: (a) testing for selection and placement, and (b) assessment of educational outcomes. Selection placement testing is based on the concept of testing developed by the French psychologist Alfred Binet, early in last century, as measurement of the ability or aptitude to profit from schooling (Broadfoot, 1996). Over the years, this has been the most pervasive role of assessment, coordinated with work and conceptions of individual differences in human intelligence. Selection testing attempts to measure students' abilities prior to a course of instruction where individuals are being properly diagnosed, placed, included, or excluded. This kind of testing later became linked with certification and exemplifies the power and control aspects of assessment as well as its role in cultural and social reproduction (Broadfoot, 1996; Gipps, 1999; Linell, 2009).

When assessing and grading students, teachers make decisions according to values, criteria, or interests (e.g. Broadfoot, 1996; Linell, 2009). People make decisions together with others because, as a system, values and interest are integrated into the regulations of established social practice. Furthermore, people accept, internalise, and act, agreeing shared ideas that they believe are true and valid, and because social practices are supported by power agreements (Broadfoot, 1996; Linell, 2009).

Features of assessment at the impact level

Assessment at the impact level is assessment to support learning and engages several fields of education associated with teachers' professional development,

school improvement, and, more generally, with educational change (Gardner, 2006).

It has long been recognised that assessment supports learning as well as measuring it (Black & Wiliam, 2003, 2006). According to these authors, Scriven and Bloom were the first to use the term ‘formative evaluation’ relating to the curriculum and teaching and the first to use the term in its currently accepted sense. They defined summative assessment tests as ‘those tests given at the end of episodes of teaching (units, courses, etc.) for the purpose of grading or certifying students, or for evaluating the effectiveness of a curriculum’ (Black & Wiliam, 2003, p. 117). Scriven and Bloom contrasted these with ‘another type of evaluation which all who are involved—student, teacher, curriculum maker—would welcome because they find it so useful in helping them improve what they wish to do, which they termed “formative evaluation”’ (p. 117, *ibid*). The concepts of formative and summative assessment are distinguished only at the time when the evaluation in question is carried out.

A variety of investigations have established that formative assessment produces learning improvements greater than those found in almost all other experiments (Gardner, 2006). Both formative assessment and assessment for learning emphasise the role of assessment supporting learning, according to ARG (Assessment Reform Group, ARG, 2002). ARG are a group of researchers who have proposed a set of changes as the basis for a new system of national assessment in the UK. The group collects and uses research from around the world to better understand how assessment can support learning:

The process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go and how best to get there (ARG, 2002) (In Gardner, 2006, p. 2).

The practice of assessment for learning follows certain principles. For example, it is part of effective planning and focuses on how students learn; it fosters motivation and promotes understanding of goals and criteria; it helps learners to know how to improve and it develops their capacity for self-assessment (ARG, 2002). These are some of the principles proposed by the Assessment Reform Group (ARG, 2002) that summarises the position assumed in the research literature and that have become widely accepted.

6.2. The concept of assessment and grading

What kind of assessment are we trying to understand in this thesis?

In this thesis, assessment focuses on teachers' planned processes of gathering and integrating information about students to discover, collect, and record students' learning as a means of making decisions about their grades. These processes provide the grounds for recording and reporting students' progress but they also constitute a mechanism for the accountability of national educational provision, of individual schools, individual teachers, and communities (Skolverket, 2005). Thus, in practice, both assessment and grading are used to make decisions about students, and to regulate and legitimate the process and outcomes of competition.

Assessment and the relationship with learning theory

The research suggests many approaches to learning and assessment and points to the essential nature of the alignment of assessment with learning, teaching, and content knowledge as a basis for claims for the validity of any assessment (James, 2008). This relationship cannot be taken for granted because there are some assessment practices that have only a vague relationship with current understandings of learning, for example, the 'short answer test' in science that requires recall of taught facts, ignores processes, does not take into account the ways that students learn the subject matter, the difficulties students confront, and the way that these issues are dealt with (ibid). Learning is viewed as the conditioned response to external stimuli and this theory of learning is referred to as 'behaviourism'. A basic belief of the behaviourist school is that learning is seen as linear and sequential. Other assessment practices may be based on the cognitive constructivist views of learning. This latter approach focuses on the individual acquisition of knowledge but recognises that learning is rarely identical to what is taught and that learning is not only a means for absorbing knowledge but an active process of individual sense-making. The implication for assessment in this view of learning is similar to the behaviouristic approach but may focus on some problem-solving and understanding of a scientific idea. The sociocultural view of learning and assessment assumes that learning occurs in an interaction between the individual and the social environment. This learning and assessment approach leads to student development and is regulated by the interaction between students and teachers, focusing on the role of feedback. The implication for teaching and assessment is that teachers need

to create an environment that enables students to think beyond their current level of competence. This view of learning will be described in greater detail below.

To develop a consistent assessment system with any degree of concordance between assessment practices and belief about learning, it is important to consider how these different approaches might be blended or synthesised. In order to draw together a variety of approaches, the research usually draws a distinction between summative and formative assessment. These approaches are often referred to as assessment of learning and assessment for learning. Both types of assessment are important and central to an effective educational practice and there is a considerable amount of research on how to use these different approaches— which are conceptually and pragmatically distinct (e.g. Harlen, 2005; Black, Harrison, Hodgson, Marshall & Serret, 2010).

Assessment in contemporary societies

Assessment is undergoing a paradigm shift from psychometric to a wider pattern of educational assessment because of changes in the social purposes of assessment in recent years (Broadfoot, 1996). Broadfoot views formal assessment as an integral part of the provision of mass education and associates it with the increasing complexity of the division of the labor that is leading to an increase in the complexity and regulation of many aspects of social life. She argues that the actual practice of assessment is operating as a series of checks and balances on the education system to guarantee its major function of perpetuating the social, economic, and political status quo. In this context, the research on assessment suggests that assessment has acquired a much broader profile and is required for a wider range of purposes, for example; assessment for learning, assessment to provide feedback to students, as well as assessment to provide information about students, teachers, and schools and to steer curricula and teaching. Thus, assessment also functions as an accountability procedure, as a tool for selection and as a certifying device. The traditional psychometric model is no longer adequate for the current circumstances; hence, the assessment paradigm has shifted.

How assessment and grading are currently used

Currently, one can ask what assessment means for students and teachers and what the teachers' priorities are when assessing students for the purposes of

grading. The concept of assessment is almost unknown to teachers and students, and often it is used whilst talking about grading because neither teachers nor students make any distinction between these processes (McMillan, Myran & Workman, 2002; Selghed, 2006; Lysne, 2006; Rinne, 2013; Gómez, 2016). Grading is also the most common concept when teachers need to communicate the results that students have achieved at school (e.g. Selghed, 2006; Rinne, 2013). In Scandinavian countries, in general, most people, including teachers and students, are unfamiliar with the term assessment but the term refers to the quantitative measuring of students' outcomes: 'most people use the word grading synonymously with assessing student progress and achievement level' (Lysne, 2006, p. 328). Lysne's view corresponds to that of Newton (2007), who also suggests that all kinds of assessment and grading lead to decisions by teachers and the school. According to Gipps (1999) and Broadfoot (1996) grading is essentially a psychometric model used to measure individual differences. These authors have described the paradigm shift from this model to a new, more educational model that involves descriptions of a wider range of evidence and feedback to support learning; this is assessment.

Further distinctions are necessary in regard to assessment. The term classroom assessment (in everyday classroom activity) is often equated with formative assessment (e.g. Stiggins, 2006; Bennett, 2011). When using the concept classroom assessment in this thesis, it does not refer to the assessment processes included in formative assessment — the kind of assessment used in the work of Black (2003) and Black and Wiliams (1998, 2006) — but it means the process of collecting learning evidence at the judgement level that provide the grounds for grading students.

Much of the confusion in terms of assessment context has to do with changed educational ideology (Lysne, 2006). According to Lysne, the term assessment is used synonymously with evaluation in spite of the fact that they have different meanings: 'Evaluation implies some degree of judgment regarding goodness, worth and values, and is more sensitive to cultural values and general purposes of education as stated in most national school legislation' (Lysne, 2006, p. 328). The concept of evaluation corresponds quite well with that put forward by Eva Forsberg and Viveca Lindberg (2010), who explain the difference as follows: 'Assessment is seen as assessments that focus on individual student achievement while evaluation is linked to the evaluation of a specific curriculum or a program' (p. 42). Thus, there are good reasons to differentiate between the terms assessment and evaluation. This study is,

however, focused on how teachers assess individual student achievement, and how they gather information and use this information for the purposes of grading, and not the evaluation of any system.

6.3. A sociocultural view of learning and assessment

The current thesis adopts a sociocultural approach to learning. This approach describes learning and knowing as a social process, situated in physical as well as sociocultural contexts (Lave & Wenger, 1991; Rogoff, 1990; Vygotsky, 1987). The context of the study — the scene — is the school, and the purpose of the study is to examine human action. Thus, the thesis is concerned with describing, interpreting, and possibly explaining actions. Wertsch (1998) emphasises that the task of a sociocultural analysis is to understand how mental functioning is related to the cultural, institutional, and historical context: ‘the task of a sociocultural approach is to explain the relationship between human action, on the one hand, and the cultural, institutional, and historical contexts in which this action occurs, on the other’ (Wertsch, 1998, p. 24). The Vygotskian heritage lies in the focus on the mediational capacity of speech. The focus of Wertsch’s idea is a basic irreducible description of agency as individual-acting-with-mediational-means. The focus of the mediated action and the cultural tools in the action make it possible to live ‘in the middle’ and to address the sociocultural situatedness of action, power, and authority (Wertsch, 1998, p. 65). Both the individual agent’s role in mediated action and the cultural tools are involved in the actions. This is because the mediational means or cultural tools are situated culturally, institutionally, and historically. Learning and thinking are always situated in a cultural setting and are dependent on the utilisation of cultural resources. Drawing on the ideas of Lev Vygotsky, Mikhail Bakhtin, and Kenneth Burke Wertsch (1998) has explored the relationship between individual and social communication, emphasising the multi-voicedness of communication and the linguistic dimensions of speech. He maintains that we can use mediated action, which he defines as the irreducible tension between active agents and cultural tools, as a productive method of explicating the complicated relationships between human action and its multiple cultural, institutional, and historical contexts.

6.4. A discursive approach

The communicative actions between teachers and students are a collaborative accomplishment. The teacher and student are interdependent in the social event of speech. It is through language, written or spoken, that people can act and interact. But teachers use a range of modes of communication that go beyond spoken language, such as gestures, diagrams, images, and demonstrations. The focus of the current thesis is the observation of classroom activities, interpreted as language and actions involving other modes of communication that teachers and students draw upon to negotiate meaning and achieve understanding. According to Fairclough (1992), language can be regarded, amongst other social structures, as defining certain potentials and possibilities and excluding others. Both teachers and students articulate particular modes of language use alongside the social relations that occur in the classroom. Thus, language is an element of social practice, controlling particular areas of social life. The way that people use language within a specific domain, such as the school, is organised in sets of patterns that Fairclough (2003) calls ‘discourse’ or ‘discursive orders’. New orders of discourse include expressions, such as ‘globalisation’, ‘modernity’, ‘neoliberalism’, etc. Neoliberalism is a political project involving a series of reforms in society, for example, the marketisation and commodification of the school. It involves new modes of being and acting, which in turn, leads to new forms of language use (Ibid). Some discourses are dominant; others are subordinated within a society at particular times. When subordinated discourses or ‘voices’ are suppressed and certain other discursive orders attain hegemony, we are dealing with monological language practices. The alternative approach to monologic language is dialogical language, or dialogicality, in this case the speech becomes contextualised and de-privileged. Fairclough has drawn on earlier ideas on language from Foucault and Bakhtin’s ‘dialogical’ theory of language and these are mostly concerned with sociocultural processes. In the dialogical theory of language, discourse is understood not as a process involving individuals and social structures as independent entities, but, instead, as dynamic and mutual interdependencies between individuals.

7. Method

7.1. The pedagogic setting, the research methods, and the analysis of data

The setting

The pedagogic setting for this study is composed of six municipal schools that offer the natural science program in different communities in the region of Skåne. The schools, teachers, and students were chosen using probability sampling (Robson, 2007); schools, teachers and students were randomly selected and asked to participate in the study. The only criterion was that the teachers were working in upper secondary schools that offered the natural science program and thus were certified and active teachers in chemistry, physics, or biology. The natural science program is one of the six national higher education preparatory programs. With a qualification from the program, students should have the knowledge needed for higher education primarily in the natural sciences, mathematics, and technology.

The overall research strategy

The empirical work of the thesis consisted of three stages and three different and separate qualitative research methods, each of which began with a pilot study to adjust and fine-tune the chosen method. Thus, multiple methods of data collection and analysis were used to address the research question of this thesis, namely: (a) individual face-to-face interviews with 25 science teachers working in five different schools at the upper secondary level (b) an ethnographic study in three different classrooms; chemistry, biology and physics in a municipal school that provides a natural science program. Every

class consisted of around 30 students each. The observations were followed by interviews with each of the three teachers involved in the ethnographic observations. The teachers and students of this second stage of the project did not participate in the empirical work of the first stage and (c) group conversations with 15 of the students from the three different science classes involved in the ethnographic observations. The students included 2 from biology class, 6 from chemistry class and 7 from physics class.

Individual face-to-face interviews with 25 teachers – Interviews 1

With the help of school catalogues edited by the schools, 40 licensed science teachers from various municipalities in Skåne were contacted via e-mail and telephone and asked to participate in the interviews. 25 teachers from five schools in three different communities of Skåne agreed to participate in the research project. The teachers were interviewed face-to-face using a semi-structured approach for 25–30 minutes in their workplaces and asked about their assessment practices. The questions were categorised into three main areas: (a) teachers' understanding of the different qualities of science education knowledge, such as procedural skills, analytical skills, and critical thinking, as well as how they assessed these qualities (b) what skills teachers considered when assessing student learning and what kind of assessment was critical when grading students and (c) the way in which students participated in the assessment process.

It was my personal preference to begin the empirical work with interviews with the teachers because the interviews gave me an overall view of teachers' assessment practices. The use of language facilitates contact with teachers as talk is a way of understanding what lies behind their actions as assessors of students' learning. Language is the principal mode of meaning-making. It mediates the communication that makes it possible to know how teachers think about their own assessment practices; their inner speech through which the teacher's thoughts are brought under conscious control. As my thesis is based on sociocultural theories of learning it is appropriate to consider language as a method for finding answers to my research questions, as inspired by both Vygotsky (1987) and Bakhtin (1986). For Vygotsky, language is the tool embedded in action and, as such, gives rise to meaning. Yet for both Vygotsky (1987) and Bakhtin (1986), meaning is dependent on the social and historical context in which it occurs. The interviews with teachers may provide a sense of how teachers themselves consider their own assessment. The task of the

interviewer is to create a conversational atmosphere with the teachers and to allow them to talk, listening more than speaking, enjoying the task and following other practicalities as recommended by Robson (2007). From a sociocultural point of view, interviews as a research method run the risk of decontextualising human action by separating action from practice, and therefore, it is of central importance to go to the place where the teachers' speech originates (Kelly & Roth, 2006). An ethnographic study enables the researcher to observe teachers in action and to map these observations onto what teachers have said about their experience as assessors during the interviews, and also what they have not said.

Ethnography

The second stage of the empirical work consisted of an ethnographic study of three different science classes in one communal school that offers different vocational programs and programs for further studies, including the science program in Skåne. An ethnographic approach is the description and understanding of the life and traditions of people living in various cultures. The method once focused on indigenous peoples but is now commonly used more generally (Robson, 2007). A full ethnography requires participation in the culture for a period of months or years, but it is also possible to conduct ethnography on a small scale, as in the present thesis. There are many variations of the ethnographic approach depending, among other things, on how much the observer participates in the action. None of the possible options are problem-free. The approach used in this thesis is termed observer-as-participant (Robson, 2007), where the observer does not participate in the observed activity but whose status as a researcher is known to the participants. This approach is opposite to that of participant-as-observer, where the observer tries to establish a relationship with members of the group during the observation. In the ethnographic approach used for this thesis, it is important to do everything possible to ensure that the observed individuals are unaware of being observed. During the first days of the observations this was very difficult to achieve, but after being in the classroom for a few days both the teachers and the students began to ignore my presence. Prior to the observations, an ethnographic pilot project was conducted in another municipal upper secondary school in order to practice all the practicalities involved in the process of observation, such as how to manage the video camera, the sound, lighting, the placement of the camera, and the additional audio devices that were used to observe the conversations of individual

students, all in accordance with Hammersley and Atkinson (2007), and Robson (2007). It is important to emphasise that when observing as a participant, all the people involved in the ethnography should know and accept the role of the observer. It was part of the observer's preparation in the very early stage of the observations to provide the students and teachers with information about the purpose of the research and the ethical considerations that were involved.

I have completed a total of 20 observations resulting in 50 hours of activities recorded without interruption in the three classes in sequential lessons over a period of four weeks. The observations were of a first-year biology course composed of 25 students who were 16 years old, a third-year physics course composed of 25 students who were 18 years old, and a first-year chemistry course composed of 25 students who were aged 16 years old. Both teachers' and students' activities during the lessons, experimental sessions, and other usual activities in the classroom were video and sound recorded. Thus, the subject matter of the study involves the careful examination of teachers' and students' conceptions of assessment in science classrooms. Group conversations consisting of two to four students were also recorded in parallel with teacher lecturing, as well as students' teamwork and experimental sessions, with the goal of capturing some students' conversations. The aim of the recorded conversations was to study the discourses that contributed to student learning and assessment across the subject matter domains. Language is the primary human mediator of student science learning and we are interested in understanding the use of this tool in the social construction of content knowledge, emphasising its use in guiding both learning and teachers' assessment. The focus of interest was to understand the systematic assessment practices employed by the teacher as well as the students' roles as learners and assessors of their own work.

The interviews with teachers and students – Interviews 2

The third stage of the empirical work entailed face-to-face interviews with the three teachers involved in the ethnographic observations and some of the students from these observed classes. Thus, this stage began during the breaktimes (e.g., between lessons or during lunch) in the school where the ethnographic observations were being made. A series of two or three cycles of interviews were conducted with every group of students, with the aim of performing a Gateway narrative inspired by Caroline Lundford Mears (2009). This kind of narrative is a series of interviews conducted with groups of

students or teachers in a form of conversation. The conversations with the group of students consisted of three cycles of conversations with every group to explain the meaning of the ethnographic study, the thesis, their rights as narrators, etc. Staging the interviews as multiple conversations allowed for communication to be improved and the learning of the individuals who agreed to participate in this project to be maximised. The interviews with different groups of students were conducted on several occasions, taking advantage of school time work breaks. Only the last interviews were recorded and the students were asked to book a special time for this last conversation due to the need to record it. A total of 15 students agreed to participate in the group interviews. Students created their own groups of two to four for every group conversation. The conversations with teachers and students that followed the ethnographic observations provided an opportunity to interpret human action. It also provided an opportunity to understand what had been observed in action in the classroom during the lessons, and the possibility to ask teachers and students about the motives for their actions. For example, teachers and students were asked about teachers' assessment practices during the interactions and why the students were usually silent during the lessons. All the empirical data collected were analysed by focusing on teachers' assessment during interaction with students, as well as the students' participation in learning and in the assessment process.

7.2. Analysis of the videos, sound recordings, and interviews with the teachers and students

Interviews with 25 teachers – Interviews 1

The interviews with teachers were structured using Robson's (2007) qualitative research design and conducted as individual face-to-face interviews with open-ended or semi-structured questions. The interviews were analysed by categories and subcategories, using two coding cycles in the Verbal Coding Exchange System (Saldaña, 2013). The data were then displayed in matrixes (Miles & Huberman, 1994) to visualise and generalise teachers' statements about their assessing and grading practices. Similar statement and relationships were arranged in clusters. This is the first step or level of the coding process, which is followed by a second level of coding and the development of subcategories.

The ethnographic research

The collection procedure for the data was performed by an observer who sat silently in three different science classrooms and audio- and video-recorded most of the activities in which students and teachers were involved over a period of four weeks. This stage of the study used an ethnographic approach where the researcher was involved solely as a researcher rather than a full-scale ethnography conducted in a short time and undertaking a degree of participation. The participation of the observer involved the observation of events, interactions, and behaviours, and the production of field notes. All observations were audio-recorded and video-recorded following Martin Hammersley and Paul Atkinson's (2007) methods in ethnography.

A variety of forms of verbal exchange were drawn from the transcripts to categorise the unit(s) as a pattern of 'skilled conversations' (Saldaña, 2013, p. 137) representing a higher or deeper level of information, including exchanges such as debates, inquiries, contradictions between teachers and students, and acts of negotiation (Gomez & Jakobsson, 2014). We also tried to find patterns of communication showing how the teachers assisted students' learning by analysing their teaching and assessing the taught material. Students' agency and engagement in learning and assessment were also carefully analysed, including how both teachers and students exerted their roles in the subject; for example, how the students engaged or did not engage in science discourse with teachers and peers, and how teachers did or did not answer students' questions, and how they assessed answers from students. The analysis also focused specifically on how students' ideas were evaluated and shared by teachers and peers, and on the way in which teachers provided opportunities for students to discuss, share, and negotiate students' ideas and interest in the subject matter.

Interviews with three teachers – Interviews 2

At the end of the observations, the three teachers involved were asked to participate in in-depth interviews. The aim of the interviews was to listen to teachers talk about what had been observed during the lessons, in order to reach an understanding of the teachers' actions in specific situations. We asked teachers for their views on particular events that had been observed during their interactions with students and about the probable causes of outcomes. The interviews were analysed in one cycle of the coding exchange system (Saldaña, 2013).

The findings are reported in article II: Everyday Classroom Assessment practices in Science classrooms in Sweden (Gomez & Jakobsson, 2014).

Interviews with students – Interviews 2

In the final stage of the overall research project, a total of 15 students from the three classes involved in the ethnographic study were interviewed. I call students narrators. Their voices are an essential factor in the reflection process. The interviews were face-to-face in the form of multiple group conversations.

When analysing gateway narratives, one must take the students' narratives seriously. I started by analysing what one of the groups said and tried to identify patterns in their narratives that matched with the other groups. I then performed a cross-case analysis to shed light on what could be learned from the other groups of narrators: what did the data mean across all the narrators? Overlaying, coinciding or contradicting narratives were analysed and considered in the subsequent findings. I examined the group narratives many times to identify patterns in the narratives, contradictions and differences among the diverse groups, and to transcribe and analyse the data for everyone that participated in the groups. In order to consider each narrative that had specific interest for each of the research questions, a separate document for each research question was developed and reviewed, highlighting patterns related to every research question. The findings are reported in article III: Students' explanations of their science teachers' assessments, grading practices and how they learn science (Gomez, 2016).

7.3. Delimitations

According to Wertsch (1998), almost all human actions are mediated actions. This thesis does not address all kinds of actions and mediational means within the studied context because it is limited to the classroom environment. Thus, it does not involve all agents that are immersed in the school context and it will not examine the tensions in mediational means in the context in any detail, neither will it consider the simultaneous goals of a mediated action. The focus of the thesis is to examine teachers' and students' activities related to teachers' assessment and grading practices. The classroom is the tool for interpreting human action and the scene, the agents in action are the teachers and students; they are observed in the ethnographic study and, a posteriori, asked about their motives for undertaking such action.

The thesis does not examine the processes of local and national evaluation of schools and teachers or school policy. The national Swedish assessment and grading system is described in this thesis a *grosso modo* to illustrate the studied context but it is not meant to examine how the complete Swedish assessment system works. I have not studied how the national curriculum works in upper secondary school after the changes in 1994 and 2011, however I have looked at some paragraphs from the curriculum, specifically those that are directly related with assessment and learning. From a pedagogical point of view, it is not possible to neglect the relationship between assessment and learning. The national curriculum describes the teachers' and students' roles and responsibilities in the processes of learning and assessment.

One basic mission of science education is to understand an idea, and to be able to use it to explain natural phenomena scientifically, being aware that the scientific explanation is only one of a multiplicity of alternatives. Both the discussion of a scientific explanation, and the dialogues on a multiplicity of alternative explanations, require students to talk about science. The same applies for students who wish to demonstrate their ability to convey critical thinking and to address, for example, ethical questions about the environment or genetic engineering. In this context, it is illustrative to describe what the natural science program looks like, to explain teachers' and students' views on learning and assessment. Thus, I have in the current research used some paragraphs from natural science programs to illustrate methods to learn science.

The current thesis considers teachers' assessment practices only in municipal schools, excluding high-performing and national recruitment programs existing in these schools, and does not take into consideration independent, adult, distance and other kinds of schools that offer natural science programs. Neither does the thesis consider issues related to age, gender, and ethnicity.

7.4. Methodological considerations

Making action the focal point of a sociocultural approach to understanding how teachers assess students' acquisition of knowledge requires paying attention to how teachers act in an authentic learning and assessment situation such as a classroom. It is in the real classroom situation that teachers' and students' roles can be observed, to understand the assessment processes that occur during interactions with students. In the first stage of this thesis, 25 teachers were interviewed with the purpose of understanding their perceptions of their own assessment practices. These interviews may partially help us to understand the conceptions that teachers have of their own assessment practices and what they perceive is going on when interacting with students in their class. In the second stage of the empirical work for this thesis, an ethnographic approach was, therefore, thought to be more likely to be fruitful in describing, interpreting, and explaining how a class of students interact, experience, and make sense of their everyday activities. According to Säljö (2000), thought and language are not equivalent and therefore one cannot know the individuals' reasons and their thoughts are not always disclosed through conversations. What people say is only a general expression for accomplishing the communicative convention for social interaction that is in force at that moment. From a sociocultural perspective, this means that during an interview it is not possible to follow teachers' and students' thoughts, but it is possible to follow what they say and what they do. From a methodological point of view, doing ethnography makes it possible to observe interactional processes in a class, for example, dialogues and negotiation processes between students and teachers as well as between students. In order to develop explanations for specific situations during the interactions in class, the ethnographic observation was complemented by interviews with teachers and students that were conducted when the observations were complete. Both formal and informal conversations with teachers and students may help to overcome some of the difficulties when

using interviews as a research method (e.g. Lemke, Kelly & Roth, 2006; Robson, 2007). The conversations allow the interviewer and interviewee to enter into a dialogical situation that may help to make sense of situations in which teachers and students' have interacted, and that may reveal something of both the teachers and the students' conceptions about assessment (e.g. Linell, 2001, 2009). Furthermore, conversations may cast light on the power dimensions of the relationships between students and teachers that might not be readily observable in the ethnographic observation.

A limitation of interviews as a research method, when they are used alone, is that they do not take language, power, and societal issues into account. This is discussed by Lemke, Kelly and Roth (2006), who suggest that what happens during an interview cannot be analysed purely on its own terms. The interview 'data' must be understood in relation to other conversations and other events in our lives, and in other situations (see also Linell, 2001, 2009). What emerges from conversations with students and teachers must therefore be seen as situated not only in their own thoughts and their own conceptions; what they express must be understood in light of what is seen as socially reasonable as well as what one is expected to say in different situations.

By combining ethnographic findings with teachers' and students' statements, and by using established assessment theories, an understanding of some of the complex processes that take place in a science classroom is enabled. As a sociocultural approach maintains, human action is involved with mediational means in a fundamental relationship, as individual(s)-acting-with-mediational-means (Wertsch, 1998). This means that any attempt to draw conclusions from the qualitative data produced in observations and interviews exclusively in their own terms, and taking only one of these three entities in account, would be misleading.

7.5. My experience and standpoint

Shortly after I began work as a chemistry teacher, I reflected and engaged in discussions about learning science, and, in parallel with my work as a teacher, I have immersed myself in this area with further studies in pedagogy at Lund University. When I commenced my doctoral studies in the field of assessment in sciences, I wondered why assessment had now become a focus of attention for researchers and the media while neither students nor parents or teachers are aware of what is really involved in the term assessment. I have been teaching

for almost my whole professional career, which began in the 1970s, and I have always assessed my students' performance. I asked myself: what is the point of this new kind of assessment I am now learning? Some answers to my question came during my very first international conference on pedagogical assessment at Stockholm University (November 2007). On this occasion, two important things happened that inspired my research and that have had important emotional and pedagogical consequences for the choice of the current research. First, I had the opportunity to learn about the different purposes of assessment from a lecture by a renowned expert in the field of assessment, the English professor Dylan Wiliam, who presented assessment as a model for classroom transactions, as something new that you, as a teacher, have the possibility to change. He meant that assessment is a value-driven procedure about students, about achievement, and about learning. He introduced me to assessment for learning (Afl). In my earlier practice as a teacher, I had placed learning at the center of teaching but underscored the importance of assessment of that learning. Thus, since the beginning of my doctoral studies, I have been interested in assessment that supports learning because I have reflected on this approach. Second, during the conference, another expert in the field of assessment, Viveca Lindberg, discussed the different ways teachers assess students. She claimed that we really do not know how teachers assess students because, in Sweden, there is no research on how teachers in fact assess students' performance in the classroom, and she argued that this is a weakness in the field of assessment in the Swedish context. Her statement made an impact on me and further inspired me to undertake the current thesis.

Given that assessment has a social role in selection, certification, and control, and because it is an important tool for regulating competition, assessment should be considered as having an epistemic and social character. In this context, my interest in science education and the assessment of students' performance in sciences has developed with the emerging perspectives of the different images of the nature of science. I have reflected on these and become particularly concerned with science as a human activity. As a chemist and a teacher, I have experienced how several forces have tended to picture natural sciences as difficult, authoritarian, and available only to people with special talents. In many situations, these conceptions of science make students feel stupid, yet, in spite of the specialised modes of reasoning and calculation in science education, are no more complex than other subjects (e.g. Lemke, 1990). Science is also commonly pictured as having an aura of absolute

objectivity and we teach our students complex and subtle skills that we expect them to use in answers to written tests.

8. Findings

8.1. The findings, the overall research project and foci of the three published articles

The aim of this thesis was to investigate science teachers' assessment and grading practices in Swedish upper secondary schools. According to the teachers' statements, we found in interviews with the 25 science teachers that assessment is performed continuously in the classroom and that the summative approach at the end-of-term involves two main operations: (a) gathering information on assessment sheets from a variety of sources, such as examinations, homework, and experiment reports and (b) combining collected information in an interpretative synthesis limited to an arithmetic algorithm that lead to grades. The test is the most important assessment tool and the most critical evaluation approach through which teachers decide on grades. The process by which teachers determine the grades takes place without participation from students. Assessment is generally considered to be an activity in which teachers alone fully exercise their autonomy. All the teachers claim that students' results on written tests are their primary source of information for determining end-of-term grades, although some teachers gave some importance to classroom activities, as well as homework and reports, especially if the teacher was not sure that the written tests provided a comprehensive picture of a student's performance in the subject matter. The teachers considered other, more qualitative, sources of information, usually as complementary information when students were in danger of failing. Teachers indicated that they assessed students' oral participation during the lessons, but it was unclear as to how and in what situations they assessed students in this way or how this influenced grading. Some teachers expressed positive thoughts about self-assessment whereas other teachers held negative views of such methods. Only one teacher worked with portfolios of each student's work and discussed the nature of sciences (NOS) with them. A further finding was that teachers do not differentiate between assessment and grading and they used

both terms as equivalent during the conversations. When the teachers were asked questions about assessment they shifted the subject and talked about grading. Most of the teachers complained about students' begging for higher grades, and claimed that grade inflation was common in their schools.

The findings of the first study of teachers' assessment practices have been published in the article, Science teachers' assessment and grading practices in Swedish upper secondary schools (referred to hereafter as article I). The conversations with teachers in this research project provide an image of the way that teachers themselves experience their own assessment practices and this was an important starting point for further investigation of teachers' assessment practices inside the classroom. This thesis stems from sociocultural approaches and considers it important to observe classrooms interaction between teachers and students in action to understand the dynamic process of assessment in the classroom and to answer the question: what do teachers actually do when they assess students for grading during classroom activities? For example, it is meaning that students develop their ability to think critically and analytically. When teachers assess and grade students, should they consider all knowledge students have acquired outside the school. The teachers claimed that all these qualities of knowledge and skills are assessed by the test. Thus, in addition to written examinations, classroom observations were undertaken in order to examine the approaches that teachers have developed for assessing critical and analytical thinking, reasoning and understanding in science, and its relation to society. The classroom observations were performed as an ethnographic study and the results have been published in the article, Everyday classroom assessment practices in science classrooms in Sweden (referred to hereafter as article II). The focus was to observe how teachers perform assessment during interactions in the classroom, and how teachers assess students' abilities, for example, to conduct a discussion, to argue and draw conclusions and analyse and solve problems. The ethnographic study has shown that in classroom practices the teachers' discourses are dominant; the I-R-E pattern of interactions characterised the lessons, but teachers did not always answer students' questions and they controlled the discussions in a way that prevented students from developing their learning. In general, students were not given the opportunity to analyse, discuss or challenge the teachers' discourses. Students spoke quietly to each other when they did not understand the subject or when they were disappointed with the teachers' answers to their questions. One of the observed teachers claimed during the individual interview that students prefer the lessons to be silent. Other teachers claimed that a silent classroom is the teacher's preference.

In article II, some of the teachers' statements, found in the interviews for article I, about their own assessment philosophies, have been confirmed by observing teachers and students in action. The classroom observations show, for example, in line with findings presented in article I, that the everyday classroom assessment is characterised by the I-R-E pattern of communication and that the written test is the most important approach to assessing students. The discourse of written tests dominated the classroom talk for both students and teachers. The written tests seem to be the most important assessment tools the teachers use for grading. This approach is used to assess all kinds of knowledge and skills. The study shows that students have very limited opportunities to discuss their own questions, in part because the lessons are dominated by the teachers' discourse. However, it was found that in addition to the written tests, the three teachers assess students through informal methods that include students' individual characteristics. Furthermore, these kinds of teacher assessments were noted in individual conversations with the teachers in conjunction with the ethnographic observations. These teachers' discourses included students' characteristics, for example, 'weak student', 'top student' and 'student that asks stupid questions'. This suggests that teachers use both formal and informal assessment approaches to make decisions about the assessment of students. One teacher claimed that she used these informal assessments to grade students.

The third and final empirical phase in this thesis was to investigate the teachers' assessment and grading practices from the students' perspectives. In this phase, students had the opportunity to explain, in small groups together with their peers, the different conceptions they have of learning and assessment sciences. An observation, made in the classrooms during the ethnographic work, was the dominance of the written test in the discourse of both teachers and students, a finding that is in line with most of the recorded material, both in the whole class and in the group interviews with the students.

A further observation in the ethnographic study was that students were mostly silent in lessons. Therefore, interviews with the student groups focused on asking them why they were silent. Teachers and students explained the silence in the classrooms in similar ways. Some teachers claimed that everybody, both teacher and students, wanted it this way. Some students confirm this by explaining that silence is expected, but some other students claimed that they were afraid to talk for fear of being labelled by their peers as stupid and, therefore, they preferred to be silent and were accustomed to this dynamic. The students' explanations of teachers' assessment have been

published in the article, Student explanations of their science teachers' assessments, grading practices and how they learn science (referred to hereafter as article III).

8.2. How the findings reported in the three published articles provide answers to the research question of the thesis

Article I, Science teachers' assessment and grading practices in Swedish upper secondary schools, highlights teachers' assessment practices in upper secondary school in accordance with the teachers own experiences from their assessment practices. The written test is the most critical method for assessing students with the purposes of grading, including assessment of different qualities of knowledge, such as procedural and analytical skills, familiarity, as well as aptitude for critical thinking. This finding answers research question 1, 'What kind of assessment approach is crucial for teachers when determining a student's grade and how do teachers assess students' knowledge in terms of procedural and analytical skills, familiarity, and aptitude for critical thinking?'. Moreover, teachers believe that they solely have responsibility for assessing and grading students, that students should not participate in the assessment process and that students should receive grades at the end of the term. Teachers do not use some of the formative assessment approaches, such as self- and peer assessment. This finding answer research question 3 'How are students involved in the assessment process and what are teachers' and students' experiences of peer- and self- assessment?'.

Article II, Everyday classroom assessment practices in science classrooms in Sweden, highlights teachers' and students' conceptions of teachers' assessment and grading in the classroom context and present impressions of how science is taught and learned, both from the teachers' and the students' perspectives. The teachers practice assessment during interaction with students through different methods, for example, through the I-R-E pattern. In the I-R-E pattern of learning, teachers control the discourses in the classroom, preventing open discussions with students about issues of interest to them or about issues pertaining to the specific subject. Teachers control the classroom discourse while students were silent, obtaining scientific information from the

teacher as passive elements in learning. This approach does not allow students to show how skillful they are in the subject, students cannot show how adept they are at arguing or exercise critical thinking in these situations. Students' own questions are not taken seriously and students make the decision to be silent to avoid ridicule from teachers and peers. This finding answers research question 2, 'To what extent and in what ways do science teachers practice assessment during classroom interactions and what do the instructional and learning discourses look like in these situations?'

Article III, Student explanations of their science teachers' assessments, grading practices and how they learn science, highlights students' perceptions of teachers' assessment and grading practices and students' conception of learning science. It was found that students do not have knowledge about how teachers assess them besides the test; that some students are irritated because teachers favoured some students when grading them; that students do not differentiate between assessment and grading. Those are some of the experiences of assessment explained by students. According to students, in science programs it is not necessary to have opinions or discuss issues during teachers' lecturing. Students make the decision to be silent during lessons because the atmosphere in classrooms is not appropriate for talking. Students suppress their own agency in learning as they remain silent. Those findings answer research question 4, 'How do students in the upper secondary school system experience teachers' assessment and learning as part of science education and what are students' experiences in terms of their own agency in learning and assessment during lessons?'. Students explain that they do not participate in the assessment process, that the test is the most important approach for teachers to assess and grade and that they do not have knowledge about self- and peer assessment, thus this finding answers research question 3, 'How are students involved in the assessment process and what are teachers' and students' experiences of peer- and self- assessment?'.

8.3. The contribution of the writers of articles I, II and III

My own contribution to article was to plan and conduct the interviews, to write the article, and to submit it for publication. My former tutor, Professor Anders Jakobsson, was involved in the process of choosing the research method, co-analysing the data obtained, and providing support during the process of writing. My contribution to article II was to plan and conduct the ethnographic study, plan and conduct the interviews with teachers and students, write the article and submit it for publication. Professor Jakobsson provided advice on the research methods, co-analysed the data and was engaged in acquiring the necessary equipment for the ethnographic observations: camera, Dictaphone, etc. I am the sole author of the third article.

8.4. Summary of the articles

Article I

Science teachers' assessment and grading practices in Swedish upper secondary schools. *Journal for Education and Training* (2015), Vol. 2, No. 2.

María del Carmen Gómez, Anders Jakobsson

This study examines science teachers' assessment and grading practices as well as student participation in the assessment process. The teachers were asked about how and when they assess students and what was crucial when grading students. I asked teachers when they considered students to have developed the following knowledge criteria: aptitude for critical thinking, analytical and practical skills, and how they assessed students regarding these skills. The article examines overall evidence-based assessment practices as reported by teachers in face-to-face interviews. The assessment and grading practices were found to be at odds with modern perspectives of assessment as well as its role in learning.

Article II

Everyday Classroom Assessment practices in Science classrooms in Sweden. *Cultural Studies of Science Education* (2014), Vol. 9, No. 4. DOI 10.1007/s11422-014-9595-y

María del Carmen Gómez, Anders Jakobsson

In this article, I examined the extent to which, and in what ways, science teachers practice assessment during classroom interactions through everyday activities in an upper secondary school in Sweden. Framing questions included: are teachers performing an integrated assessment of students' skills as the national curriculum mandates? If so, what do the instructional discourses look like in these situations and what are students' experiences regarding their agency in learning and assessment? Teacher-led lessons in three science classrooms were video-recorded and analysed by combining ethnographic and discourse methods of analysis. Students' conversations were also recorded. We found that traditional assessment methods, such as tests, examinations, and assignments were the most common methods used to assess and grade students' learning. Teachers mostly ignored students' questions and did not provide opportunities for students to discuss or argue scientific issues, as the national science curriculum stipulates. Different aspects of knowledge stipulated in the national Swedish curriculum, such as lifelong learning, stimulation of students' creativity, curiosity, as well as their wish to explore and convert new ideas into action, and find solutions to problems, were restricted by teachers' discourses.

Article III

Students' explanations of their science teachers' assessments, grading practices and how they learn science. *Cultural Studies of Science Education* (2016), Vol 11, No 3. DOI :10.1007/s11422-016-9740-x.

María del Carmen Gómez

This study draws on data generated through group interviews with students who were involved in a larger ethnographic research project conducted in three science classes with the purpose of understanding teachers' assessment practices in an upper secondary school in Sweden. In these interviews, I asked students about their conceptions of what forms of assessment were prioritised by their teachers, why students were silent during teachers' lectures, and the

students' experiences of peer- and self-assessment. The research design and analysis of the findings derives from what students told me about their teachers' assessment and their experiences of learning science. Students told me that, in addition to the results of the written test, they did not know what else teachers assessed and used to determine grading. Furthermore, students did not participate in classroom discussions because of peer-pressure and fear of disappointing their peers, or of being graded less favourably as a result of their views, answers, or questions. Students' silence is also linked with students' conceptions of science learning and their experiences of teachers' traditional methodologies of teaching and learning science.

9. Discussion

This thesis offers an opportunity to understand how science teachers practice assessment from both the teachers' and the students' perspectives. Although the research was undertaken in only six municipal schools which offer natural sciences education in southern Sweden, and the number of participants was limited to 28 teachers and 15 students, the findings have the potential to establish a dialogue with different agents in other contexts. Through ethnographic observations and interviews with teachers and students, this research allowed both parties directly involved in assessment and learning to offer their opinions and experiences. Explanations from teachers and students about science teachers' assessments have, to some extent, provided the opportunity to understand science teachers' assessment practices in those environments.

The teachers' perspectives on assessment provide an opportunity to review how they experience their own assessment practices. The current research found a culture among teachers that weakens and damages the foundation of a meritocratic society. This is expressed, for example, in the informal characterisation of students as 'low achieving student', 'top student', 'student who asks stupid questions', and such-like, characterisations that are based mainly on a student's characteristics or performance. It is reasonable to expect, for example, a 'top student' to achieve well and be awarded high grades. I suggest that this is a way students are labelled by teachers who believe that ability and performance are constant and changeless and, perhaps even innate and very easy to assess. Thus, the label assigned will add some bias to the teachers' observations and will, most likely, also translate into bias in the grading of the student. Such cultural phenomena are well known to be present, but are difficult to prove in specific cases. However, the mere existence of such bias decreases the accountability of grades and undermines the foundation of a meritocratic society, that is, one that is based on equality between individuals and fairness in assigning merit. In a meritocratic society, true and correct validation of merit is crucial. From this perspective, the grades obtained by any student must be both reliable and valid. This, in turn, implies that the

assessment and judgement, on which the grade is based, must also be both reliable and valid. The standard method for obtaining this is to base the assessment and judgment on verified foundations that are known and recognised by all agents with an interest in the resulting grade. This provides accountability for grades as the criteria on which they are based is transparent. Research on these issues should benefit students and society.

During the ethnographic study, I reflected on the reasons why students were silent during the teacher-led lessons and how students perceive their own silence in relation to learning the science content. I suggest that students were not regarded as co-constructors of scientific knowledge or of classroom learning settings, in part because their questions and thoughts were not always answered (Gomez & Jakobsson, 2014), and in part because they were silent during the lessons; and therefore, passive elements in learning. Thus, in this thesis, I have described how the teachers deal with students when they ask questions and how the students experience their agency in the processes of learning and assessment. Teachers want students to participate in the science discourse but expect them to do so in accordance with the ‘social order’ (Fairclough, 2001) established in the classroom. Teachers assume a position established from the beginning of the term depending on their own subject position; to be heard, and to not listen to students. This is possible because it is in line with their discursive rights and obligations. Both teachers and students are allowed, and required, to speak or not speak within particular situations. The observed classroom interaction is typical of teacher-centred instruction or the I-R-E approach. The results of this thesis match the findings of previous research. For example, Rop (2003) found that under teacher-centred instruction, classroom questions remain unanswered. The possibilities for enriching dialogue based on the remaining uncertainties in the minds of students are closed. In this respect, I recall a quote from an observed classroom in the ethnographic study where a student posed a question to the teacher, the teacher answered:

Do not get yourself entangled with transition elements!!! (...) they behave silly, I do not want to talk about them!! They are anomalous! Forget them ... until you begin to study chemistry at university!! (Gomez & Jakobsson, 2014, p. 33).

This way of responding to students may be a consequence of teacher-centered instruction where teachers establish the social order. This finding is in line with the research on science education, which shows that science classrooms tend to discourage dialogue because teachers are more focused on fulfilling the curriculum and ‘doing school’ (Rop, 2003, p. 13). Moreover, teachers do not

encourage questioning because of the view that ‘to get higher grades students do not need higher thinking’ (ibid, p. 23). Thus, the research and the findings in this thesis suggest that there is an effort to teach content, whilst ignoring deeper learning and understanding. This is a persistent problem in science education stemming from the dominance of traditional learning that focuses on concepts. In this approach, science is not considered as contextualised and may not reflect the position of individuals in society.

The silence in the observed classroom and reported in Gomez & Jakobsson (2014) may also be interpreted as a lack of communication or non-participation, that is, a lack of action that may have implications in learning and the assessment of students’ performance. If we return to Fairclough (2001), teachers talk and learners do not talk as observed in classroom unless they are directed to do so by the teacher. In this sense, the silence seems to have a symbolic connotation as students identified their silence as a form of agency, as a strategy for avoiding ridicule from their peers (peer pressure). Thus, silence occurs not only because of cultural characteristics and teaching methodology. Social inclusion and students’ agency appear to be important factors in the students’ silence. To deal with this, teachers need to develop instruction methodologies to encourage students to develop skills and to understand when students are conscious of being silent, to involve and include them in classroom activities. Students also fear that teachers will ridicule them, or they fear that their speech may negatively influence teachers’ assessments. Thus, silence is a form of decision-making. A clear example of silence or non-participation is when students and teachers claim that even if students know the answer to the teacher’s question, they do not talk. Thus, science teachers need to develop an understanding of silence and all its possible causes and recognise that any decision or action is grounded in agency. It is useful to pay attention to the research on the role of emotions in the process of learning (e.g. Järvelä, 2011) because the silence of students may be also related to the emotional learning climate in the classroom, where peer ridicule or the fear of looking stupid can lead to non-active participation through silence. Since teachers have control over classroom interactions, they are important agents in establishing and normalising pattern of interactions in classroom discourse and may possibly consider silence to be appropriate behaviour. More research on this issue is needed.

The teaching of content at the expense of understanding has also been confirmed in the present study. Both teachers and students claimed that the subject of physics is not contributing to dialogue or leading students to engage

in critical thinking. A student in physics claimed: ‘It depends on the subject. If the subject requires that you must have opinions—for example social studies and history—then you must talk. It is different in physics’ (Gómez, 2016). It may be possible that when science is taught without consideration of students’ opinions this may contribute to silence among students. The student in this example has the perception that science, specifically physics, has no need for discussion. This results in a silent student. This position on students’ agency is related to the finding that, in the culture of science classrooms, students feel that they may be punished if they talk.

The pattern of perceptions found amongst teachers and students regarding science learning and assessment may be problematic. The history of the development of scientific ideas shows that the expansion of knowledge involves critical thinking, dialogue, and questions, and that school is supposed to be the place to introduce students to the world of scientific thinking. The approach of learning science content by ‘doing school’, as described by Rop (2003), was found among teachers in the current thesis. One teacher, who also refused to give students the opportunity to participate in the assessment process, claimed:

They [students] are not involved in the assessment process at all (...). It is me who should possess the knowledge (...) it is me who knows what it is the truth, the right things (...) it is me who is the assessor! (Joachim). (Gomez & Jakobsson, 2015, p. 15).

It was found that teachers do not explain to students how they assess them for the purposes of grading other than through tests. This contradicts one of the principle goals in science education, the need for science teachers to develop assessment instruments that are contextualised and go above and beyond paper-and-pencil experiences. Other than in tests, students do not know what teachers assess and they associate assessment directly with grades:

Patrik: Ah ... Assessment ... We don’t really know. We have only got grades for the tests we have done and it is the grades you obtain in the test that count (...) what we have right because of our results on the written test and assignments (Gomez, 2016, p. 15).

These assessment practices are in accordance with a Swedish review of research on assessment that observed that the psychometric tradition dominates teaching and assessment from comprehensive school to higher education (Lindberg, 2005). Teachers’ assessment is not an authentic measure of science knowledge because current practices do not support the development of a

critical scientific literacy. Paper-and-pencil tests are even less effective to make scientific knowledge socially relevant.

Conducting integrated assessments of students' knowledge and skills is a complex and crucial part of instruction which involves new ways of being and acting and should respond to students' needs rather than discouraging them. Teachers need to become important transformational agents of assessment, and practice a form of assessment which follows the meritocratic ideal. But teachers should not be exposed to the strong pressure, from various agents, of assigning ever higher grades, as teachers and students have revealed in the current research. According to the interviewed teachers, the pressure for higher grades comes from principals and students, as the teacher Regina claimed when she was asked about the way students participate in the assessment process:

(...) Most students accept their grades because they have got the grades they expect, but there are some students who try to influence me as they need higher grades in order to be admitted to the physician education program. Can I get higher grades, they say? No, I say! Regina (Gomez & Jakobsson, 2015, p. 13).

The issue of higher grades was perceived in different ways from the perspective of students. There was a claim from some students that teachers showed favouritism towards particular classmates when grading them, and that some students used tactics to obtain higher grades, as a student claimed when she was asked about her experience of self-assessment:

There are many peers which are good at talking about their grades, they are good at marketing themselves (Gomez, 2016, p. 22).

The findings of this thesis describe how teacher assessment and grading practices do not develop into a fully effective learning environment. The requirement to, as completely as possible, assess many different aspects of students' knowledge and skills, like aptitude for scientific methods and critical thinking, requires a broad, largely continuous and sometimes informal process of assessment, incorporating also classroom activity. Under the influence of teachers' assessment practices, individuals in the learning environment are not encouraged to learn to their full potential. Assessment by grading has been shown to exclude students, since it is not fully transparent and is influenced by some factors irrelevant to grading. In view of the problems identified with teachers' practices of assessment by grading, as reflected by teachers and students in this thesis, it is reasonable to assume that grades are not fully

reliable and valid for natural science in upper secondary Swedish schools and that students do not use their learning potential in full.

When the current grading system was imposed, it lacked, to a major extent, clear grade criteria. Also, any tool to calibrate grades was, and formally still is, lacking. The system completely relied on the professionalism of the teachers to grade their students, a situation which still prevail. Simultaneously there are a number factors in force, like competition between schools and evaluation based on grades, which adds pressure to award higher grades. Since this pressure for higher grades is not counteracted by any controlling mechanism, it is likely that the sum of all grades nationally will increase, irrespective of teachers striving to keep their judgements consistent over time. In this view, the current general 'grade inflation' in Sweden is not surprising but is an obvious sign of a lack of consistency in grading. The current inconsistencies in grading have led the authorities to pursue a more extensive use of national tests as a tool for calibrating grades (e. g. Waldow, 2014).

If grades are not assigned in accordance with a governing, democratically agreed, clear, and usable set of instructions, variations in the criteria for the same grades may arise. The resulting differences in the content of grades may undermine trust in the grades as fair and just, and threaten the base of the meritocratic ideal. If, as in Sweden today, unreliable or invalid grades are used in the evaluation of institutions, distribution of educational resources, control of admittance to higher education then the decisions based on these grades will lack rational grounds, since the grades on which the decisions are based do not reflect any specific content. Consequently, a reliable and valid system of grading is crucial in many aspects. This may be obtained by a well calibrated grading system, and a well performed, consistent system of assessment by grading. It is evident that the current Swedish grading system has flaws, and that the grading process deserves to be a field for future research.

References

- Aikenhead, G. S. (2006). *Science Education for Everyday Life. Evidence-Based Practice*. New York & London: Teachers College Press, Columbia University.
- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House*, Vol. 78, No. 5, 218–223.
- Anderson, K. T., Zuiker, S. J., Taasoobshizazi, G. & Hickey, D. T. (2007). Classroom discourse as a tool to enhance formative assessment and practice in science. *International Journal of Science Education*, Vol. 29, No. 14, 1721–1744.
- Ash, D., Crain, R., Brandt, C., Loomis, M., Wheaton, M. & Bennett, C. (2007). Talk, tools, and tensions: Observing biological talk over time. *International Journal of Science Education*, Vol. 29, No. 12, 1581–1602.
- Bakhtin, M. M. (1986). *Speech Genres and Other Late Essays*. Y. McGee, trans. Austin: University of Texas Press.
- Beach, D. (1999) Om demokrati, reproduction och förnyelse I dagens gymnasieskola. *Pedagogisk Forskning i Sverige*, Årgång 4, Nr 4, 349-365.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, Vol. 18, No. 1, 5–25.
- Black, P. (2003). The importance of everyday assessment. In: Myron Atkin, J. & Coffey, J. (Eds.). *Everyday Assessment in the Science Classroom: Science Educators' Essay Collection*, p. 1-12.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, Vol. 5, No. 1, 7–73.
- Black, P. & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, Vol. 29, No. 5, 623–38.
- Black, P. & Wiliam, D. (2006). Developing a theory of formative assessment, in *Assessment and Learning* edited by John Gardner. London: SAGE Publications, p. 81-100.
- Black, P., Harrison, C., Hodgen, J., Marshall, B. & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice*, Vol. 17, No. 2, 215–232
- Broadfoot, P. M. (1996). *Education, Assessment and Society*. Buckingham & Philadelphia: Open University Press.

- Broadfoot, P. M. & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education*, Vol. 11, No. 1, 17–27.
- Broekkamp, H. & Van Hout-Wolters, B. H. A. M. (2007). Students' adaptation of study strategies when preparing for classroom tests. *Educational Psychology Review*, Vol. 19, No. 4, 401–428.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, Vol. 30, No. 2, 123–142.
- Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106, 429–458.
- Canal, L., Bonini, N., Micciolo, R. & Tentori, K. (2012). Consistency in teachers' judgements. *European Journal of Psychology of Education*, Vol. 27, No. 3, 319–327.
- Carter, L. (2007). Sociocultural Influences on Science Education: Innovation for Contemporary Times. *Science Education*. 92(1), 165–181.
- Cazden, C. B. (2001). *Classroom Discourse. The Language of Teaching and Learning*. Portsmouth, NH: Heinemann.
- Corrigan, D., Dillon, J. & Gunstone, R. (2007). *The Re-Emergence of Values in Science Education*. Rotterdam: SENSE Publishers.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, Vol. 58, 438–481.
- Dahlstedt, M. (2007). I val(o)frihetens spår: Segregation, differentiering och två decennier av skolreformer. *Pedagogisk Forskning I Sverige*, Vol. 12, No. 1, 20–38.
- Daniels, H. (2008). *Vygotsky and Research*. London & New York: Routledge.
- Davis, A. (1998). *The Limits of Educational Assessment*. Oxford: Blackwell Publisher.
- Duit, R. (2007). Science education research internationally: Conceptions, research methods, domains of research. *Eurasia Journal of Mathematics, Science & Technology Education*, Vol. 3, No. 1, 3–15.
- Ecclestone, K. & Pryor, J. (2003). 'Learning careers' or 'assessment careers'? The impact of assessment systems on learning. *British Educational Research Journal*, Vol. 29, No. 4, 471–487.
- Elwood, J. (2006). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education*, Vol. 13, No. 2, 215–232.
- Engström, S. & Carlhed, C. (2014). Different habitus: Different strategies in teaching physics? Relationships between teachers' social, economic and cultural capital and strategies in teaching physics in upper secondary school. *Cultural Studies of Science Education*, Vol. 9, No. 3, 699–728.
- European Commission Community Research (2007). Union européenne. Commission européenne, & Union européenne. Direction générale de la recherche (2007). *Science education now: A renewed pedagogy for the*

- future of Europe. Office for Official Publications of the European Communities.
- European Commission (2011). Science Education in Europe: National policies, Practices and Research.
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge: Polity Press.
- Fairclough, N. (2001). *Language and Power* (2nd Ed.). London: Pearson Education Limited.
- Fairclough, N. (2003). *Analysing Discourse. Textual Analysis for Social Research*. London & New York: Routledge.
- Ford, M. J. & Wargo, B. M. (2012). Dialogic framing of scientific content for conceptual and epistemic understanding. *Science Education*, Vol. 96, No. 3, 369–391.
- Forsberg, E. & Lindberg, V. (2010). *Svensk forskning om bedömning – En kartläggning*. Vetenskapsrådet. Vetenskapsrådet rapportserie, 2:2010. Stockholm.
- Gardner, J. (2006). *Assessment and Learning*. Edited by John Gardner. London: SAGE Publications.
- Gee, J. P. (2003). Opportunity to learn: A language-based perspective on assessment. *Assessment in Education*, 10(1), 27-46.
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355–392.
- Glaser, R. & Silver, E. (1994). Assessment, testing, and instruction: Retrospect and prospect. *Review of Research in Education*, Vol. 20, No. 1, 393–419.
- Gómez, M. (2016). Students explanations of their science teachers' assessment, grading practices and how they learn Science. *Cultural Studies of Science Education*, 3(11). DOI:10.1007/s11422-016-9740-x.
- Gómez, M. & Jakobsson, A. (2014). Everyday classroom assessment practices in science classrooms in Sweden, *Cultural Studies of Science Education*, 9(4), 825-853. DOI:10.1007/s11422-014-9595-y.
- Gómez, M. & Jakobsson, A. (2015). Science teachers' assessment and grading practices in Swedish upper secondary schools. *Journal for Education and Training*, 2(2), 1-19.
- Gott, R. & Duggan, S. (2002). Problems with the Assessment of Performance in Practical Science: Which Way Now? *Cambridge Journal of Education*, 32(2), 183-201.
- Greeno, J., Moore, J. & Smith, D. (1993). Transfer of situated learning, In: D. Detterman and R. Stenberg (eds) *Transfer on Trial: Intelligence, Cognition and Instruction*. Nordwood, N. J.: Ablex Publishing Corporation.
- Hammersley, M. & Atkinson, P. (2007). *Ethnography: Principles in Practice*. Third Edition. London & New York: Routledge.
- Hargreaves, E. (2007). The validity of collaborative assessment for learning, *Assessment in Education*, Vol.14, No. 2, 185–199.

- Harlen, W. (1999). Purposes and procedures for assessing science process skills. *Assessment in Education*, Vol. 6, No. 1, 129–144.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning — tensions and synergies. *The Curriculum Journal*, Vol. 16, No. 2, 207–23.
- Harris, L. R. & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, Vol. 16, No. 3, 365–381.
- Hayward, E. L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, Vol. 22, No. 1, 27–43.
- Herrenkohl, L. R. & Guerra, M. R. (1998). Participant structures, scientific discourse, and student engagement in fourth grade. *Cognition and Instruction*, Vol. 16, No. 4, 431–473.
- Hodson, D. (2009). *Teaching and Learning about Science: Language, Theories, Methods, History, Traditions and Values*. Rotterdam; Boston: Sense Publishers.
- Hsu, P. & Roth, W. (2014). From authoritative discourse to internally persuasive discourse: Discursive evolution in teaching and learning the language of science. *Cultural Studies of Science Education*, 93(3), 729–753.
- Huot, B., O'Neill, P. & Moore, S. (2010). Usable past for writing assessment. *College English*, Vol. 7, No. 5, 495–517.
- Huston, T. (2005). Race and Gender Bias in Student Evaluations of Teaching. Research report. Retrieved from http://www.seattleu.edu/cetl/cetl_006.htm
- Jakobsson, A., Mäkitalo, Å. & Säljö, R. (2009). Conceptions of knowledge in research on students' understanding of the greenhouse effect: Methodological positions and their consequences for representations of knowing. *Science Education*, 93(6), 978–995. DOI: 10.1002/scs.20341.
- James, M. (2008). Assessment and learning. In: Swaffield, S. (Ed.) *Unlocking Assessment. Understanding for Reflection and Application* (p. 20–35). Abingdon, Oxon: Routledge.
- Järvelä, S. (Ed.) (2011). *Social and Emotional Aspects of Learning*. Oxford, UK: Elsevier Academic Press.
- Johnson, D. (2003). Activity theory, mediated action and literacy: Assessing how children make meaning in multiple modes. *Assessment in Education*, Vol. 10, No. 1, 123–129.
- Kane, M., Crooks, T. & Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, Vol. 18, No. 2, 5–17.
- 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*, Vol. 14, No. 2, 181–199.

- Klapp Lekholm, A. & Clifordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation: An International Journal on Theory and Practice*, Vol. 15, No. 1, 1–23.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, MA: Cambridge University Press.
- Lemke, J. (1990). *Talking Science. Language, Learning and Values*. Westport, CT and London: Ablex Publishing Corporation.
- Lemke, J., Kelly, G. & Roth, W.-M. (2006). Forum: Toward a phenomenology of interviews. Lesson from the phenomenology of interviews. *Cultural Studies of Science Education*, 1(1), 83–106.
- Leung, C. & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: Assessment as discourse and assessment of discourse. *Language Testing*, Vol. 21, No. 3, 335–359.
- Lidström, L., Holm, A.-S. & Lundström, U. (2014). Maximising opportunity and minimising risk? Young people's upper secondary school choices in Swedish quasi-markets. *Young*, Vol. 22, No. 1, 1–20.
- Lindberg, V. (2005). Svensk forskning om bedömning och betyg 1990–2005 [Swedish research on assessment and grading 1990–2005]. *Studies in Educational Policy and Educational Philosophy*, 2005(1). Stockholm: School of Teachers Education, Stockholm University.
- Linell, P. (2001). *Approaching Dialogue*. Amsterdam: J. Benjamins Pub.
- Linell, P. (2009). Rethinking language, mind and world dialogically. Interactional and contextual theories of human sense-making. In: Valsiner, J. (Ed.) *A Volume in Advances in Cultural Psychology: Constructing Human Development*. Charlotte, N.C.: Information Age Publishing, INC.
- Lundahl, C. & Folke-Fichtelius, M. (2010). *Bedömning i och av skolan- praktik, principer, politik*. Lund: Studentlitteratur AB.
- Lundahl, C. (2010). Nationella prov – ett redskap med tvetydiga syften. [National test – a tool with ambiguous purposes]. In Lundahl, Christian & Folke-Fichtelius, Maria (red.). *Bedömning i och av skolan- praktik, principer, politik*. P. 223-240. Lund: Studentlitteratur AB.
- Lundford Mears, C. (2009). *Interviewing for education and social science research. The gateway approach*. London: Palgrave Macmillan.
- Lundqvist, E., Almqvist, J. & Östman, L. (2012). Institutional Traditions in Teachers' manners of teaching. *Cultural Studies of Science Education*, 7:111–127. DOI 10.1007/s11422-011-9375-x.
- Lysne, A. (2006) Assessment theory and practice of students' outcomes in the Nordic countries. *Scandinavian Journal Educational Research*, Vol. 50, No. 3, 327–359.
- McMillan, J. H., Myran, S. & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4), 203-213.

- Matthews, M. R. (2012). Changing the focus: From nature of science (NOS) to features of science (FOS). In: Khine, Myint Swe (Ed.) *Advances in Nature of Science Research*, 3-26.
- Mehan, H. (1979). *Learning Lessons, Social Organization in the Classroom*. Cambridge, MA & London: Harvard University Press.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, Vol. 23, No. 2, 13–23.
- Miles, M. B. & Huberman, A. M. (1994). *An expanded sourcebook. Qualitative data analysis. Second Edition*. Oaks, CA: SAGE Publications.
- Mintzes, J. J., Wandersee, J. H. & Novak, J. D. (2001). Assessing understanding in biology. *Journal of Biological Education*, Vol. 35, No. 3, 118–124.
- Mortimer, E. F. & Scott, P. (2003). *Meaning Making in Secondary Science Classrooms*. Philadelphia: Open University Press, p. 3-26.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, Vol. 14, No. 2, 149–170.
- Nyström, E. (2007). *Talking and Taking Positions. An Encounter Between Action Research and the Gendered and Racialised Discourses of School Science*. Umeå, Sweden: Umeå University.
- O’Byrne, B. (2009). Knowing More than Words Can Say: Using multimodal assessment tools to excavate and construct knowledge about wolves. *International Journal of Science Education*, 31(4), 523-539.
- Ollin, R. (2008). Silent pedagogy and rethinking classroom practice: structuring teaching through silence rather than talk. *Cambridge Journal of Education*, Vol. 38, No. 2, 265–280.
- Osborne, J., Simon, S. & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049-1079.
- Osborne, J. & Dillon, J. (2008). *Science Education in Europe: Critical Reflections. A report to the Nuffield Foundation*. London: Nuffield Foundation.
- Osborne, J. & Hennessy, S. (2003). Literature Review in Science Education and the Role of ICT: Promise, Problems and Future Directions. *FUTURELAB SERIES*.
- Osborne, J., Simon, S. & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049-1079.
- OECD (2009). *Equally Prepared for Life? How 15-Years-Old Boys and Girls Perform in School*. Paris: OECD.
- OECD (2012). *Closing the Gender Gap: Act Now*. OECD Publishing.
- OECD. (2013). *Gleichstellung Der Geschlechter: Zeit Zu Handeln*. Paris: OECD.

- Pedretti, E. (2002). T. Kuhn meets T. Rex: Critical conversations and new directions in science centres and science museums. *Studies in Science Education*, 37, 1–42.
- Pellegrino, J. W., Chudowsky, N. & Glaser, R. (2001). *Knowing what Students Know: The Science and Design of Educational Assessment*. USA: The National Academy of Sciences.
- Poehner, M. E. (2011). Dynamic assessment: Fairness through the prism of mediation. *Assessment in Education: Principles, Policy & Practice*, Vol. 18, No. 2, 99–112.
- Poehner, M. E. & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, Vol. 9, No. 3, 233–265.
- Radford, J., Ireson, J. & Mahon, M. (2006). Triadic dialogue in oral communication tasks: What are the implications for language learning? *Language and Education*, Vol. 20, No. 3, 191–210.
- Rasmussen, A. & Friche, N. (2011). Roles of assessment in secondary education: Participant Perspectives. *Educational Assessment, Evaluation and Accountability*, 23, 113–129.
- Rinne, I. (2013). Att bedöma och bemöta — Förståelsen av betyg så som den visar sig i ett betygssamtal på gymnasiet. *Nordic Studies in Education*, 33(03), 171–186.
- Roberts, R. & Gott, R. (2006). Assessment of performances in practical science and pupil attributes. *Assessment in Education*, Vol. 13, No. 1, 45–67.
- Robson, C. (2007). *Real World Research*, (2nd Edition). Oxford: Blackwell Publishing.
- Rogoff, B. (1990) *Apprenticeship in Thinking: Cognitive Development in Social Context*. New York: Oxford University Press.
- Rop, C. (2002). The meaning of students' inquiry questions: A teacher's beliefs and responses. *International Journal Science Education*, Vol. 24, No. 7, 717–736.
- Rop, C. (2003). Spontaneous inquiry questions in high school chemistry classrooms: Perceptions of a group of motivated learners. *International Journal of Science Education*, Vol. 25, No. 1, 13–33.
- Sadler, T. D. (2009). Socioscientific issues in science education: Labels, reasoning, and transfer. *Cultural Studies of Science Education*, Vol. 4, No. 3, 697–703.
- Saldaña, J. (2013). *The coding manual for qualitative researchers*. Second Edition. Los Angeles: Sage.
- Säljö, R. (2000). *Lärande I Praktiken. Ett socioculturellt perspektiv*. Norstedts akademiska Förlag.
- Sampson, V. & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and

- recommendations for future directions. *Science Education*, Vol. 92, No. 3, 447–472.
- Scott, P. H., Mortimer, E. F. & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, Vol. 90, No. 4, 605–631.
- Selghed, B. (2006). *Betygen I Skolan — Kunskapssyn, Bedömningsprinciper och Lärarpraxis*. Stockholm. Liber.
- Simms, M. & George, B. (2014). Approaching assessment from a learning perspective: Elevating assessment beyond technique. *Educational Assessment, Evaluation and Accountability*, Vol. 26, No. 1, 95–104.
- Skolverket (2005). *National Assessment and Grading in the Swedish School System*. Stockholm. [The Swedish National Agency for Education].
- Skolverket (2007). *Provbetyg-Slutbetyg-Likvärdig bedömning? En statistisk analys av sambandet mellan nationella prov och slutbetyg i grundskolan, 1998–2006*. [The Swedish National Agency for Education]
- Skolverket (2009). *Likvärdig betygssättning i gymnasieskolan? En analys av sambandet mellan nationella prov och kursbetyg*. [The Swedish National Agency for Education]. [Equivalent grading in secondary schools? An analysis of the relationship between the national sample and Course Rating].
- Skolverket (2011). *Gymnasieskola 2011*. Skolverket. Stockholm. [The Swedish National Agency for Education]
- Skolverket (2011). *Nationella prov i gymnasial utbildning*. [National tests in upper secondary schooling]. Stockholm: Skolverket. [The Swedish National Agency for Education].
- Skolverket (2012). *Curriculum for the Upper Secondary School. Natural Science Programme, 2012-10-31*. Stockholm: Skolverket. [The Swedish National Agency for Education].
- Skolverket (2013). *Curriculum for the Upper Secondary School*. Stockholm. [The Swedish National Agency for Education].
- Smith, B. P. & Hawkins, B. (2011). Examining student evaluations of black college faculty: Does race matter? *The Journal of Negro Education*, 80(2), 149-162.
- SOU (2008). *Statens Offentliga Utredningar, 2008:27*.
- Stiggins, R. J. (2006). Assessment for learning: A key to motivation and achievement. *Edge*, Vol. 2, No. 2, 3–19.
- Swaffield, S. (Ed.) (2008). *Unlocking Assessment. Understanding for Reflection and Application*. Abingdon, Oxon: Routledge.
- Treagust, D. F. & Duit, R. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*, Vol. 3, No. 2, 297–328.

- Tytler, R. (2007) Re-imagining Science Education — Engaging students in science for Australia's Future. ACER Australian Council for Educational Research.
- Vygotsky, L. S. (1987). *Thinking and Speech*. Edited by RW Rieber & AS, translated by N. Minick. Translation of Russian original [1934]. (cf. Vygotsky, 1986). New York & London.
- Waldow, F. (2014). Conceptions of Justice in the Examination Systems of England, Germany, and Sweden: A Look at Safeguards of Fair Procedure and Possibilities of Appeal. *Comparative Education Review*, 58(2), 322-343. doi:10.1086/674781.
- Wells, G. & Mejía Arauz, R. (2006). Dialogue in the classroom. *The Journal of the Learning Sciences*, Vol. 15, No. 3, 379–428.
- Wertsch, J. V. (1998). *Mind as Action*. Oxford University Press.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, Vol. 7, No. 9, 703–713.
- Wikström, C. (2005). Grade stability in a criterion-referenced grading system: the Swedish example. *Assessment in Education*, 12(2), 125-144.
- Wikström, C., & Wikström, M. (2005). Grade inflation and school competition: An empirical analysis based on the Swedish upper secondary schools. *Economics of Education Review*, 24, 309–322.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.
- Wolf, D., Bixby, J., Glenn III, J. & Gardner, H. (1991). To use their minds well. *Review of Research in Education*, Vol. 17, No. 1, 31–74.
- Wyatt-Smith, C., Klenowski, V. & Gunn, S. (2010). The centrality of teachers' judgment practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, Vol. 17, No. 1, 50–75.
- Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon P. R., & Furtak, E. M. (2008). On the Impact of Formative Assessment on Students Motivation, Achievement and Conceptual Change. *Applied Measurement in Education*, 21, 335-359.



Science Teachers' Assessment and Grading Practices in Swedish Upper Secondary Schools

Maria del Carmen Gómez (Corresponding author)
School of Science Education, Malmö University
Valdemarsväg 3, 224 74 Lund, Sweden
E-mail: Maricarmen7879@gmail.com

Anders Jakobsson
School of Science Education, Malmö University
Hjälmarsväg 10 Malmö 205 06 Malmö Sweden
E-mail: Anders.jakobsson@mah.se

Received: January 16, 2015 Accepted: January 29, 2015 Published: February 15, 2015

doi:10.5296/jet.v2i2.7107

URL: <http://dx.doi.org/10.5296/jet.v2i2.7107>

Abstract

This study examines science teachers' assessment and grading practices as well as student participation in the assessment process in the upper secondary school. The teachers were asked about how and when they assess students and what was crucial when grading students. We asked when they considered students to have developed the following knowledge criteria: aptitude for critical thinking, analytical and practical skills and how they assessed students regarding these skills. We report overall evidence-based assessment practices from the teachers' comments in face-to-face interviews. Teachers' comments are closely aligned and associated with long-established beliefs. The assessment and grading practices were found to be at odds with modern perspectives of assessment as well as its role in learning.

Keywords: Grades; Assessment Practices; Science Education; Upper Secondary School; Knowledge

1. Introduction

Our study is based on issues related to developments in Swedish teachers' assessment practices following school reforms in 1994 and the revised curriculum of 2011 (SOU, 2008, p. 27). In these documents, the epistemology of concepts on knowledge, understanding, confidence in the subject matter, aptitude for critical thinking and students' participation in their own assessment have been particularly emphasised. Changes to the national curriculum in Sweden and recent research in western countries suggest that assessment and grading both

require new approaches to measuring and evaluating students' learning. Our study seeks to determine how these issues relate in a Swedish context and is one of several reporting the findings of a larger research project that aims to examine teachers' assessment practices in upper-secondary science in Sweden.

International and Swedish research literature shows that teachers' assessment practices have yet to be examined (McMillan, Myran, & Workman, 2002). Further, Johnston, Afflerbach and Weiss (1993) noted the lack of research in this area and, in later studies, McMillan et al. (2002) discusses the same problem. In the Swedish context it is important to study how teachers assess and grade students as we have changed the curriculum and we need to know how teachers take into account the new guidelines on assessment. The prominent assessment research, both international and from a Swedish perspective, has above all focussed on compulsory schools (e.g., Brookhart, 1994, 1997, 2004; Brown, 2004; Cizek, Fitzgerald, & Rachor, 1995, 1996; Klapp Lekholm & Cliffordson, 2009; Martínez, Stecher, & Borko, 2009; McMillan et al., 2002; McMillan & Nash, 2000) or on the effects of classroom assessments (e.g., Alkharusi, 2008; Harlen & Crick, 2003). Teachers' literacy as regards assessment has also been reported (e.g., Brookhart, 2011; DeLuka & Klinger, 2010; Howley, Howley, Henning, Gillam, & Weade, 2013), as well as students' conceptions of assessment processes (e.g., Andersson, 2000; Brookhart & Bronowicz, 2003; Brown & Hirschfeld, 2008). Only few studies highlight factors that influence teachers' assessment practices (e.g., Martínez et al., 2009). Additionally, research exploring how teachers assess student achievement in upper secondary schools within the framework of science education is relatively rare. The current study addresses these issues by examining teachers' own statements about assessment processes. The national curriculum in Sweden prescribes students' participation in assessment; for this reason, we will additionally focus on how and in what ways they are involved. This also implies examining the extent to which teachers use features of formative assessment to serve the social construction of scientific knowledge according to the national curriculum.

2. The Research on Teachers' Assessment

The predominant research on compulsory schooling from both international and Swedish perspectives suggests that student assessment is left to individual teachers. Additionally, most research in the area shows that it is not specifically studied or understood (e.g., Brookhart, 1994; James & Pedder, 2006). However, recent studies indicate that both achievement and non-achievement are taken into account when teachers assign grades (Klapp Lekholm & Cliffordson, 2008; 2009). For example, social backgrounds, motivation, gender and ethnicity may affect grades and assessment (Klapp Lekholm & Cliffordson, 2008; 2009). These researchers suggest that teachers often try to maximise students' grade outcomes to benefit both students and schools (see also Cliffordson, 2004a).

Selghed's 2006 exploration of teachers' conceptions of the grading process in upper secondary schools showed rather disparate ideas about the Swedish assessment system. Selghed concluded that teachers graded similarly to how they did before the assessment reforms of 1994. Both Swedish and international research point to a shift from pencil-and-paper and single-response tests towards performance-based assessment in science education (e.g. Bell &

Cowie, 2001; Jakobsson, Mäkitalo, & Säljö, 2009; James & Pedder, 2006; Treagust, Jacobowitz, Gallagher, & Parker, 2001). For example, James and Pedder (2006) argued that the integration of formative assessment in science education may improve results and raise achievement standards. Another example is Treagust et al.'s 2001 study that focussed on a broader programme of teaching and assessment by following a physics class in which the teachers successively improved both effectiveness and learning requirements by encouraging students to discuss and develop their ideas and their scientific language. Another example of the shift is research concerning student perceptions of scientific concepts (greenhouse effect and global warming; Jakobsson et al., 2009). The authors found that the students were able to express their knowledge in a more developed manner and use different knowledge forms when they interacted with others and with cultural tools compared to when they were tested by paper-and-pencil tests. The authors further argued that in studies based on constructivist learning theory, students appeared to have many misconceptions regarding concepts, but their performance improved when they were allowed to discuss and interact with others.

3. Assessment in Science Education

As mentioned, several science education scholars have called for reforms concerning assessment of and for learning (e.g., Corrigan, Dillon, & Gunstone, 2007; Duschl & Osborne, 2000; Sampson & Clark, 2008; Tierney, 2006), arguing that the assumptions underlying current assessment approaches fulfilled outmoded functions. Several studies indicate that assessment procedures that focus on elementary knowledge in science often are biased as they rely disproportionately on a narrow range of skills, such as memorisation (e.g., Gallagher, 2007; Gott & Duggan, 2002; Osborne & Hennessy, 2003; Roberts & Gott, 2006). According to these scholars, paper-and-pencil testing might suffice for science programmes that only are designed to acquire facts on a subject. However, promoting more engaging activities, such as problem solving, argumentation and process skills, will require more demanding assessments. For example, Gallagher (2007) argued that these kinds of practices involve tasks that are student-led, have a societal context and may involve broader skills, such as analytical thinking, communication, critical thinking and problem solving. Such activities may be difficult to simply *measure* through paper-and-pencil examinations that require epistemological discussions about the subject. Recent research has increasingly focussed on formative processes used by teachers to master learning. Stiggins (2006) argued that assessment in modern societies 'must support the learning of all students so all can succeed at meeting standards' (p. 2). Effective classroom assessment may lead to profound achievement; therefore, it must describe students' current status completely. These arguments are in line with this article's general aim vis-à-vis science assessment processes in the Swedish curriculum. However, Black and Wiliam (2003) asserted that the development of formative assessment depends on new practices and assessment tools. They additionally stressed that research about these issues has to include 'the perceptions and beliefs of teachers about learning, about the "abilities" and prospects of their students, and about their roles as assessors' (Black & Wiliam, 1998, p. 51).

4. Summative and Formative Assessment

Findings related to the limits of traditional educational assessments (e.g., Elwood, 2006;

Pellegrino, Chudowsky, & Glaser, 2001), as well as the increasing amount of evidence over the past decade on the pedagogic potential of formative assessment (e.g., Bennett, 2011; Black & Wiliam, 2009; Poehner & Lantolf, 2005; Lundahl, 2011), have increased the public demand for school science reforms. However, according to Black and Wiliam (2003), the terms *formative* and *summative* did not apply to the assessments themselves, but rather to the functions they served. They argued that assessment development requires new methods and items that align formative and summative work. This implies that teachers' formative work would not be undermined by summative pressures because of accountability. In comparison, summative requirements might be better served by taking full advantage of improvements in teachers' assessment work. Bennett (2011) defined formative assessment as 'a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes'. He simplifies his statement as follows: 'as long as the results are used to change instruction, any instruments may be used formatively, regardless of its originally intended purpose' (ibid. p. 6).

Newton (2007) asserted that assessment for formative purposes has different accountability characteristics that fundamentally differ from summative assessments. For example, one important aspect of formative assessment is the *dialectical relationship* between teacher-peer-learner, which can be contextualised as consisting of five 'key strategies' (Black & Wiliam, 2009, p. 8).

Further, as the responsibility for learning falls on both teachers and learners, peer- and self-assessment is emphasised. However, these forms of assessment seem to be controversial. According to Wiliam (2000), self-assessment opponents often deny student objectivity, although their assertion applies primarily to summative assessment. In other words, accuracy in formative assessment above all is an issue of secondary significance, as the focus is on whether self-assessment can enhance the learning process. According to the authors, other formative features focus on the assessment agent. Traditionally it is the teacher who collects the evidence of learning and decides. However, formative assessment also includes peers and individual learners in the making of such decisions. In this respect, teachers consider students as important instructional resources.

5. Theoretical Foundations of Learning and Assessment in Research

In a classic article, Gipps (1994) discussed problems identified with most traditional assessment models, namely: *decomposability* and *decontextualisation*. Decomposability often assumes it is possible to divide complex competency learning into smaller parts, which in turn can be assessed through individual stimulus-response connections (Gipps, 1994). Gipps further argued that assessing separate skills may foster teaching practices where learning can be seen as linear and sequential, and where complex understanding occurs only when the basic constraint on learning is mastered. In decontextualisation, 'each component of a complex skill is fixed, and will take the same form no matter where it is used' (Resnick & Resnick, 1992, p. 43). However, according to the authors, teachers cannot teach a skill component in one setting and expect it to be automatically applicable and assessable in another. In addition, Gipps (1994) argued that situations of 'scaffolding' in learning processes, offered by people who may be more competent,

may be extended to assessment *in situ*. To circumvent these issues, Brookhart (2011) suggested that assessment may involve formative approaches, that is, to use assessments based on language, dialogues and collaborative developmental methodologies (Brookhart, 2011; Ash et al., 2007; Poehner, 2011; Poehner & Lantolf, 2005). These studies view learning and assessment as a productively attached process where the individuals are engaged interactively.

6. The Concepts of Knowledge and Assessment From a Swedish Perspective

Teachers' assessment practices are of increasing interest to the educational assessment society in connection with the introduction of new curricula in 1994 and a revised version in 2011. The Swedish curriculum incorporates statements to the effect that students' complete performance, including understanding, aptitude for critical thinking skills and confidence in the subject, should be assessed in order to produce a final grade. In addition, when teachers grade their students, they are to take into account *'...all information about students' knowledge in relation to the demands in the syllabus including such knowledge the students acquire by different ways, to make a comprehensive assessment of the students' knowledge during the whole course'* (Swedish Ministry of Education, 1994, p. 35). In many respects, this is a major difference from the earlier, centrally controlled system (Lgy, 1970; SKOLFS, 1992, p. 6, 24), in which the total annual grades were nationally balanced in accordance with the normal distribution; the grades in different geographic areas were designated guided by results from annual, centralised, national tests (e.g. Cliffordsson, 2008).

Regarding knowledge and learning, the national curriculum (Skolverket, 2013) stipulates:

The school's task of imparting knowledge presupposes an active discussion about concepts of knowledge, about what knowledge is important today, what will be important in the future, and also about how learning and the acquisition of knowledge take place. (p. 6).

Most of the character of the national curriculum of 1994 is confirmed in the newly revised curriculum from 2011 (SOU, 2008, p. 27). For example, the revised curriculum for the Swedish upper secondary school (Skolverket, 2013) states:

The national school system is based on democratic foundations. The Education Act (2010, p. 800) stipulates that education in the school system aims at students acquiring and developing knowledge and values. It should promote the development and learning of students, and a lifelong desire to learn (Skolverket, 2013, p. 4).

The all-round development of students, scientific ways of thinking and the ability to think critically are also stipulated as follows:

Students should develop their ability to think critically, examine facts and relationships, and appreciate the consequences of different alternatives. By these means students will come closer to scientific ways of thinking and working (Skolverket, 2013, p. 5).

Regarding assessment, the curriculum does not stipulate directly the kind of assessment teachers should practice but leaves the teachers the possibility of working with formative characteristics of assessment.

6.1 What types of knowledge and skills/competencies does Swedish Science Education require?

Besides facts and understanding about the chosen knowledge area the programme of natural science involves:

... the subjects biology, physics and chemistry together with the subject of mathematics are the core of the Natural Science Programme. (...) The education should develop students' knowledge about context in nature, about the conditions for life, about physical phenomena and events, and about chemical processes. (...). The education should stimulate students' curiosity and creativity, and their ability to think analytically. Students should develop a scientific approach.

Regarding critical thinking the Swedish Science Education requires:

Ability to think critically, reason logically, solve problems, and make systematic observations. Students should be given the opportunity to develop the ability to distinguish between statements based on scientific and non-scientific grounds. (...).

Regarding sociocultural issues for the development of society the Swedish Science Education requires:

The education should contain a perspective from the history of ideas, which means that the ideas and theories of the sciences are studied as parts of a historical process. The education should give an understanding of how science and the development of society both affect and are affected by each other and in particular highlight the role of science in questions concerning sustainable development. Students should also be given the opportunity to take part in ethical discussions of the role of science in society (Skolverket, 2012, p. 228).

7. The Study and the Research Questions

The overarching purpose is to explore whether teachers are using practices that comply with the national curricula on the concept of knowledge and student assessment participation. Thus, teachers were asked about their assessment practices in three main areas: a) understanding different qualities of science education knowledge, such as procedural skills, analytical skills and critical thinking, as well as how they assessed those qualities; b) what they considered when grading student learning and c) the way in which students participated in the assessment process.

The current study thus addresses the following research questions:

How do science teachers assess and grade students' knowledge in terms of procedural and analytical skills, familiarity and accumulated experience and aptitude for critical thinking?

What should teachers take into account when grading students?

How are students involved in the assessment process?

In what situations do teachers assess students and how is it made clear to the students that they are being assessed?

8. Methods and Analytic Procedures

The sample in this study was composed of 25 teachers in different communities in southern Sweden who were interviewed about their assessment practices. They were chosen by *probability sampling* (Robson 2007, p. 261), that is, they were randomly selected and offered to participate in the study. One important criterion was that the teachers should be science teachers in first-, second- and third-course programmes in five different schools and certified to teach chemistry, biology and physics at the upper secondary school level. The school samples are representative of those in southern Sweden responsible for upper secondary science programmes. However, as the study strives to collect qualitative data concerning the teachers' own experiences about assessment and grading processes in upper secondary schools, the aim was not to give a simple and generalisable image concerning all Swedish teachers, and instead was to describe and analyse teachers' own experiences about assessment and grading processes from a self-constructed narrative. According to Saldaña (2013), narratives of this kind communicate a category of knowledge that portrays human experiences in a way where actions and events may contribute positively and negatively to the investigation. Thus, the analysis of the teachers' narratives is considered exploratory and the aim is to create trustworthy data collection close to teachers' experiences and ideas. With this in mind, we argue that qualitative methods are more suited to our approach.

The interviews were carefully prepared and lasted around half an hour, excluding time for self-instruction and preparation (for details, see Silverman, 2010). The interview methods were inspired and structured from the perspective of Qualitative Research Interviews designed by Robson (2007), and conducted as individual *face-to-face* interviews in the form of open-ended or semi-structured questions. This implies that both the interviewer and interviewees were allowed to clear up misunderstandings and interviewees urged to expand on their responses. The interviews were audiotaped with the consent of the respondents and the teachers were asked about their experience about assessment practices in three main areas (see appendice). The names of the teachers are fictitious, and the five schools in different communities in the south of Sweden are named A to E.

8.1 Data Analysis

Our data analysis is based on and inspired by assessment research in the international science education community. In the first phase, we analysed the interviews and found categories and subcategories using two coding cycles in the Verbal Coding Exchange System (Saldaña, 2013). The data was then displayed in matrixes (Miles & Huberman, 1994) to generate all statements about the responder's own experience of assessing and grading processes. Further, in order to find patterns in the material, we coded the data by gathering similar teacher statements and relationships in small experience clusters. This step constituted the first level of the coding process, with the second level being a development of sub-categories by

examining the teachers' statements, followed by a revision of the first level. In this phase, we also reorganised and reanalysed the data that was coded during the first cycle method.

Our categorisation and sub-categorisation were inspired by Gott and Duggan (2002) and Osborne and Hennessy (2003) in that their results indicate assessment modes that focus on low-level conceptual knowledge. Osborne and Hennessy (2003) additionally suggested that more engaging science (e.g., problem-solving, argumentation and process skills) will require developing more rigorous assessment practices. Many of these suggestions involve tasks that are student-led and classroom-contextualised, and may involve skills such as analytic thinking and communication and practical observation. In this sense, our categorisation considered the teachers' statements and related them to different kinds of assessment practices in the literature.

Three categories were described in the following: Category number 1 – 'The procedures used by teachers to assess and grade students' learning' – aimed to answer the research question: How do science teachers assess and grade students' knowledge in terms of procedural and analytical skills, familiarity and accumulated experience and aptitude for critical thinking? In this first category, we found three different patterns of teachers' assessment practices, which in turn were arranged into three sub-categories:

Sub-category 1a. Process-focussed assessment

Sub-category 1b. Outcome-focussed assessment

Sub-category 1c. Implicit or inconsistent assessment strategies

In subcategory 1a, the assessment of students' development and their understanding of the subject are in focus. Teachers in this category consider the assessment process as a part of students' learning, with the written test not being the critical or dominating grade feature. The process gives teachers opportunities to discuss the social character and the nature of science.

In sub-category 1b, the teachers are more concerned with students' learning outcomes and accomplishing the curriculum. The teachers assess outcomes basically by written tests with the quantities of knowledge and curriculum compliance being decisive. The written examination is critical when assigning the final grades. In sub-category 1c, the inconsistency of the assessment is related to the sense that teachers display contradictory assessment and grades strategies. These teachers do not consider the assessment process and the test as critical for final grade assignment.

Category 2, '*What is the most critical factor for teachers when deciding on the final grade?*', aims to relate to the research question: What is critical for teachers when grading students? In this category, two main patterns were found and identified with the following sub-categories:

Sub-category 2a. Alternative assessment methods

Sub-category 2b. Traditional assessment methods

'Alternative method' refers to assessments that use examination methods not exclusively for the purposes of grading and often set aside unilateral paper-and-pencil examinations. The

teachers' assessment in this group is complex and includes different types and levels of student knowledge and skills. Teachers use several methodologies for assessment, such as portfolios, argumentation, dialogues or other assessment methods used for learning. With 'traditional assessment methods', we refer to assessment and grading practices mainly based on paper-and-pencil examinations. In this group, teachers exclude other types of assessment methods.

Category 3, *Student participation in the assessment process*, is related to the research questions about students' involvement in the assessment process and in what kind of situations teachers assess students, as well as how is it made explicit to them that they are being assessed. In this category, four different patterns were identified; accordingly, two sub-categories were built:

Sub-category 3a: Teachers assert that assessment is their responsibility. Students do not participate.

Sub-category 3b: Teachers have positive insights into students' participation in the assessment process.

9. Results

9.1 The Procedures Used by Teachers to Assess and Grade Students' Learning

The findings from our study are presented in this section, enlarging on each of the three major categories with associated sub-categories summarised above, while giving examples of teachers' statements in every category based on the interviews. The first category concerns *The procedures used by teachers to assess and grade students' learning* and relates to the research question about how the teachers assess and grade students in terms of procedural and analytical skills, familiarity and accumulated experience and aptitude for critical thinking. The three general patterns of procedures that were found in the analysis are described in the following three sub-categories.

9.1.1 Sub-Category 1a. Process-Focussed Assessment

Only two of the 25 teachers met the criteria for this sub-category. Their assessment methods mainly focus on students' learning and development processes. According to these teachers, assessment is complex, demands to be systematic and achieves its aims by a diversity of approaches. Students' performances are often assessed by evaluating their portfolios and their development and use of scientific language. Students' experimental work, argumentation skills and collaborative performance constitute important parts of the assessment process. One of the teachers addressed the topic as follows:

When they feel [...] when the students are not afraid to come into the laboratory session anymore, but they feel delight and excitement: when they want to begin the work; when they discuss with each other and with me about what chemistry is about, what happens and why; when they observe a chemical reaction; when the students feel that they have developed skills and knowledge [...] they master and manage the chemistry talk! And then [...] I know that they are ready [for the] next step. (Jens).

This statement may suggest that the teacher has created a kind of interactive tool to assess students' knowledge. The teacher pays attention to the emotional engagement of students in their learning processes (Järvelä, 2011), and claims that the students '*are not afraid to come into the laboratory session anymore, but they feel delight and excitement*'. In this respect, this part of the assessment seems to be spontaneous and to consider moments of contingency in the process of learning. Accordingly, the teacher notes the students' impact on his teaching. He further clarifies his ideas about assessment processes thusly:

I assess the students all the time, in the classroom, in the experimental session, and I write every assessment moment in my notebook that students have possibility to discuss with me. Students also have portfolios that I evaluate constantly. All this assessment is considered in the final grades. (Jens).

9.1.2. Sub-Category 1b. Outcome-Focussed Assessment

The majority (18) of the respondents were evaluated to meet the criteria in the sub-category of outcome-focussed assessment. The teachers claimed that written tests do not constitute the only foundation to assess students' development and understanding of the subject content. However, they described written tests as the most important and decisive element for grading. Teacher assessment is above all focussed on students' outcomes; explicit aims in the curriculum and the grading processes exclusively depend on written documents and measurements of knowledge. Procedural skills such as those demonstrated during experimental work are not an important factor for grading students.

One teacher in this category was asked about assessing student understanding in physics:

I don't assess classroom situations at all, I don't do it. It is what they do in the written test which counts. (Mathew).

Another teacher was also asked about student understanding and confidence in the subject:

The written test is the most important criterion for grading students. We try to define different tasks, both easy and difficult ones, so that the students can show if they are developing an understanding, skills and confidence. [...]. You can see how far the students have developed by such a discussion, how they're reasoning, if they're right or wrong. This is what gives me insight into their knowledge, understanding, confidence and all of those things. (Regina).

This teacher uses different tools, such as laboratory reports, homework tasks and language skills, as sources for the assessment process; however, these activities do not seem to be given any weight in the total assessment that leads to grades. In this context, she said the following:

Then you can look at the laboratory reports [...], the lessons and all other activities [...]. But you can't raise the grade because of the laboratory reports or something else (...). You cannot do it. It is still knowledge, both in width and depth, which is crucial when grading the students. (Regina).

Another teacher answered a question about assessing critical thinking in physics as follows:

In physics? [5] Yes [...] they [...] they [...] I don't know, but they are very quick to learn the application of models by using the formulas, which are simplifications of reality in the world [9]. The simplification always matches reality perfectly. Physics is not the same as in social science. There you must call in [to] question [...]. Here, in physics, we are working with models that are tested [...]. In physics nothing needs to be criticised. (Michael).

Only two teachers expressed explicit aims for students needing critical thinking in science. One of these was a biology teacher that endorsed critical aptitude as relevant during discussions on ethical issues of genetic studies. However, the two teachers did not report how they assessed these skills; instead, the teachers expressed the following:

We used to construct different levels of knowledge in the test and in this way we measure all of this... confidence, familiarity and understanding. (Robert).

It is the test that decide all of this, confidence, analysis qualities of knowledge and so on. (Maurice).

9.1.3 Sub-Category 1c. Implicit or Inconsistent Assessment Strategies

About one-fifth of the teachers' assessment and grading processes were characterised by implicit or inconsistent strategies. In addition, some of these teachers expressed an uncertainty or even a contradictory view about assessment. One example of this was the following:

There is so much different information about this [what to assess] [...] there is a vast amount of information to base assessments on, that I do not feel insecure when I assign grades. (Jon).

However, when he was asked about the most decisive factor in grading, he replied with a contradictory statement:

It is the written test which is the most critical in assigning grades. I do not assess the student at all in classroom situations, I do not.... It is how they perform in the written test which accounts for their final grade. (Jon).

9.2 *The Most Critical Factor for Teachers to Take Into Account for the Final Grade*

The second category relates to what teachers take into account when grading students. In this category, two main patterns of practices were found and identified with two sub-categories:

9.2.1. Sub Category 2a. Teachers Use Alternative Assessment Methods

Alternative assessment methods are those that shift away from pencil-and-paper and single-response tests. These could, for example, be performance-based assessments, assessments of communicative skills and evaluation of students' familiarity with the subject through dialogues. In addition, the teachers in this category strive to assess students' performance using several different tools. The written test seems not to be the only critical

tool for assessment and grading. Only few teachers (2) were found in this sub-category. One of the teachers expressed the most critical factor in chemistry by stating the following:

The most critical factor is that they can communicate chemistry to me and to each other. (Jens).

Another teacher answered the same question as follows:

What is critical to me is that students mastered the scientific language. It is only by mastering the language of science [that] students can talk about new terms, new knowledge, this is important when assessing students. They must discuss and communicate with the others. (Jackeline).

9.2.2 Sub-Category 2b. Teachers that Use Traditional Assessment Methods, that is, Written Examinations

Nearly all of the teachers (23) made statements that indicated that they especially used traditional assessments methods and that the results from written tests constituted the main tool in the grading process. Mathew expressed the most critical grading factor as the following:

Grading a student simply means checking how they performed in four or five written tests during the term. It is the essential thing about it. The written test is very important for the final grade. (Mathew).

Another example is the physics teacher, Joan:

The most important [factor] for assessment and grading is the test. (Joan).

9.3 How Do Students Participate in the Assessment Process?

The third category is connected to students' involvement in the assessment process. Additionally this category aims to describe in what kind of situations the teachers assess students and how it is made clear to students that teachers are assessing them. In the analytic phase, we found two explicit patterns of students' participation that we present in the following sub-categories. The categories were built on the teachers' statements and depended on two criteria: a) teachers' ideas about permitting student participation in the assessment process; and b) teachers' statements about students' awareness of situations when teachers were assessing them.

9.3.1. Sub Category 3a: Teachers Assert that Assessment is Their Task

A majority of the teachers (20) expressed statements that fell into this category. In general, the teachers expressed that they listen to their students' opinions about different kinds of assessment activities and that they may have opinions about the grading situation at the end of term when they receive their final grades. However, there were no statements in this category that indicated any pervading student involvement during the assessment phase. Some examples of statements in this category were the biology teacher Joachim and the chemistry teacher Charles:

They [students] are not involved in the assessment process at all [...]. It is me who should possess the knowledge [...] it is me who knows what it is the true, the right things [...] it is me who is the assessor! (Joachim).

They do not assess each other or themselves or such... I do not do anything like that. I am the teacher, I am the assessor [...]. Students' duty is to perform. (Charles).

Another teacher was asked in what ways the students are involved in the assessment and grading processes:

No, [...] they do not participate in the assessment process. They cannot do it, as they are not (...), they try to influence when we discuss grading. (...). Most students accept their grades because they have got the grades they expect, but there are some students who try to influence me as they need higher grades in order to be admitted to the physician education programme. Can I get a higher grade, they say? No, I say! (Regina).

The teacher's statement may suggest that Regina does not separate between the concepts and the process of assessment and grading. She stated specifically:

they do not participate in the assessment process. They cannot do it, as they are not (...), but they try to influence when we discuss grading.

9.3.2 Sub Category 3b: Teachers Have Insights Into Students' Participation in the Assessment Process and They are Positively Open to Alternative Assessment Practices

Only three teachers expressed that they allowed students to participate in assessment. In addition, these teachers also give students self-assessment activities or adopt alternative assessment methodologies (for example, portfolios). The teachers expressed curiosity and positive ideas about alternative assessment methods that involve students. One example is the chemistry teacher, Johanna:

Oh yes, I think that we teachers should work more to improve this (...), sometimes I ask students to make self-assessment in the final of the term, but sadly this practice does not lead to anything. (Johanna).

9.4 Teacher Comments About Students' Awareness About Assessment Processes

In order to understand whether students participated in some way in the assessment process, the teacher was asked about students' knowledge of the precise moment they were assessed.

I am not sure if they know, I assess them automatically. (Johnny).

Another one of the teachers addressed the issue more indirectly. She was concerned about students' silence in the classroom and encouraged them to take part in discussions:

Often they are too conscious about assessment. It makes it difficult for them to participate in talking situations during the lessons. I usually encourage students to participate; they refuse to talk [...]. Nobody wants to talk; very often, they observe each other, specially my high achievers; they don't want to say anything. (Brigitte).

This quotation reveals a possible contradiction between the teacher's assessment for learning purposes and assessment for summative or grading purposes. The teacher is engaged in students' learning, but she probably assesses them by evaluating the answers from discussions. In this situation, the teacher's effort to encourage students to talk does not lead to participation; on the contrary, the students become silent. However, more research is needed about what assessment situations may hinder development.

10. Discussion

A summary of results indicates that most of the participating teachers in this study seem to work with rather traditional assessment and grading strategies and methods. Furthermore, analyses suggest that these are often associated with a view of knowledge that is mainly related to a memorisation of subject content and assessed in summative approaches. In addition, the results suggest that summative paper-and-pencil tests are the dominating instrument for the purpose of grading, and that assessment of students' knowledge is principally considered the teacher's task. The students do not participate in the assessment process and only have possibilities to know the result of the assessment made by teachers. In other words, the results of this study imply an explicit distinction between statements from upper secondary science teachers' assessment and grading practices and the Swedish national curriculum that stipulates students' participation in the assessment process. An example is that a vast majority of the teachers meet the criteria in the sub-category of outcome-focussed assessment. The teachers claim that written tests do not constitute the only foundation to assess student understanding of the subject content. However, they describe written tests as the most decisive element for grading. Analyses also indicate that only few teachers use performance-based assessments of communicative skills and evaluation of students' familiarity with the subject through dialogues and development of science vocabulary.

International assessment research (e.g., Brookhart, 2004; 2011; DeLuka & Klinger, 2010) stresses strategies that focus students' learning and develop their knowledge of the subject. The comments in this study generally speak for themselves, but the various findings in terms of prevalence or level of consensus show that after 20 years of changes in the national curriculum, teachers continue working with traditional assessment practices in the science classroom. We consider that methodologies used to assess the character of knowledge through mainly memorising for examinations may exclude students from articulating their thoughts and discussing scientific and societal issues. As a result, there is a risk that the development of students' critical thinking as well as more sophisticated skills such as curiosity, creativity and ability to think analytically, which the national curriculum stipulates, may be adversely affected (e.g., Corrigan et al., 2007; Duschl & Osborne, 2000; Sampson & Clark, 2008; Tierney, 2006). Additionally, Duschl and Osborne (2000) argued that students may not engage in more advanced kinds of science because the type of knowledge the teacher is assessing simply does not demand it. Poehner (2011) asserted that students might be afraid to express their thoughts or hypotheses if they feel that an incorrect answer could adversely affect their grades, or if they ask a question that exposes any lack of knowledge.

One of the teachers tries to solve the problem through allowing students to make mistakes

when discussing scientific issues and to create assessment-free moments. However, the consensus among the teachers in this study as regards assessment and grading may impact students' development of the skills that the science curriculum requires. The teachers' assessment strategies and methods seem to restrict opportunities for discussing issues concerning the role of science in society and other specific competences. The assessment environment becomes characterised by the written test culture which seems to have a decisive influence on how to organise the subject content and what knowledge forms to emphasise.

Black and Wiliam (2009) argued that if teachers focus merely on the written test, students adapt to this environment and also focus on the written test and the complex qualities of knowledge risk being disregarded. However, increased teacher awareness about the assessment environment may contribute to a better understanding of these issues and their impact on student learning. According to Brookhart (2004; 2011), teachers need support as well as explicit tools to develop as assessors and in order to implement the curriculum. An effective integration of formative and summative assessment is needed to promote students' learning, and to activate them as owners of their own learning processes (Brookhart, 2011).

11. Conclusion

Teachers continue to use traditional learning and assessment strategies in spite of the changes in the Swedish curriculum; we suggest that one reason for this may be that teachers encounter difficulties in benefitting from the research on educational assessment; further research is needed to explore this claim. Regarding the assessment environment, and following the ideas of Brookhart (2004; 2011), teachers need to create formative assessment environments and need support and instruction in doing this; research in this field has yet to be conducted.

Appendix: Symbols in Transcript Excerpts

[...] Denotes micro-pause.

[7] Denotes 7-second pause.

Underlining denotes that the word is accented or emphatic.

! Exclamation mark denotes stress or animated tone.

References

- Alkharusi, H. (2008). Effects of classroom assessment practices on students' achievement goals. *Educational Assessment*, 13, 243–266.
- Andersson, P. (2000). Att studera och bli bedömd. Empiriska och teoretiska perspektiv på gymnasie- och vuxenstuderandes sätt att erfara utbildning och bedömningar. *Filosofiska fakulteten Linköping universitet* No 68. [To study and to be assessed. Empirical and theoretical perspective in upper-secondary and adult school ways of experiencing education and assessments. *Linköping Studies in Education and Psychology* No 68].
- Ash, D., Crain, R., Brandt, C., Loomis, M., Wheaton, M., & Bennett, C. (2007). Talk, tools, and tensions: Observing biological talk over time. *International Journal of Science Education*,

29(12), 1581–1602. <http://dx.doi.org/10.1080/09500690701494118>

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85(5), 536–553. <http://dx.doi.org/10.1002/sce.1022>

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 16(3), 269–290. <http://dx.doi.org/10.1080/0969594X.2010.513678>.

Black, P., & Wiliam, D. (2003). ‘In praise of educational research’: Formative assessment. *British Educational Research Journal*, 29(5), 623–38.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–68. <http://dx.doi.org/10.1080/0969595980050102>

Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.

Brookhart, S. M. (1994). Teachers’ grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279–301.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10, 161–80.

Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106, 429–458.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.

Brookhart, S. M., & Bronowicz, D. L. (2003). ‘I don’t like writing. It makes my fingers hurt’: Students talk about their classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 10, 221–242.

Brown, G. T. L. (2004). Teachers’ conceptions on assessment: Implications for policy and professional development. *Assessment in Education*, 11(3), 301–318.

Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students’ conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3–17.

Cizek, G., Fitzgerald, S., & Rachor, R. (1995/1996). Teachers’ assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159–179.

Cliffordsson, C. (2008). Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment*, 13, 56–75.

Corrigan, D., Dillon, J., & Gunstone, R. (2007). *The re-emergence of values in science education*. SENSE Publishers.

DeLuka, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in

teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438.

Duschl, R. A., & Osborne, J. (2000). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38(2002), 39–72. <http://dx.doi.org/10.1080/03057260208560187>.

Elwood, J. (2006). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education*, 13(2), 215–232.

Gallagher, J. J. (2007). Teaching science for understanding – A practical guide for middle and high school teachers. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Gipps, C. (1994). Beyond testing – Toward a theory of educational assessment. London: The Falmer Press.

Gott, R., & Duggan, S. (2002). Problems with the assessment of performance in practical science: Which way now? *Cambridge Journal of Education*, 32(2), 183–201. <http://dx.doi.org/10.1080/03057640220147540>.

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

Howley, M. D., Howley, A., Henning, J. E., Gillam, M. B., & Weade, G. (2013). Intersecting domains of assessment knowledge: School typologies based on interviews with secondary teachers. *Educational Assessment*, 18, 26–48.

Jakobsson, A., Mäkitalo, Å., & Säljö, R. (2009). Conceptions of knowledge in research on students' understanding of the greenhouse effect: Methodological positions and their consequences for representations of knowing. *Science Education*, 93(6), 978–995. <http://dx.doi.org/10.1002/sce.20341>.

James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109–138.

Järvelä, S. (ed.) (2011). *Social and emotional aspects of learning*. Oxford, UK: Elsevier Academic Press.

Johnston, P. H., Afflerbach, P., & Weiss, P. B. (1993). Teachers' assessment of the teaching and learning of literacy. *Educational Assessment*, 1(2), 91–17.

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.

Klapp Lekholm, A., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 15(1), 1–23.

Lgy (1970). 70 Läroplaner for gymnasieskola [Curriculum 1970 for Upper Secondary

School].

Martínez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgments, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment, 14*, 78–102.

McMillan, J. H., & Nash, S. (2000). Teachers' classroom assessment and grading practices. Decision making. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research, 95*(4), 203–213.

Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook. Qualitative data analysis* (2nd ed.). Oaks California, SAGE Publications.

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education, 14*(2), 149–170.

Osborne, J., & Hennessy, S. (2003). Literature review. In *Science education and the role of ICT: Promise, problems and future directions*. London: Futurelab Series. Report 6.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. The National Academy of Sciences, USA.

Poehner, M. E. (2011). Dynamic assessment: Fairness through the prism of mediation. *Assessment in Education: Principles, Policy & Practice, 18*(2), 99–112.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233–265.

Resnick, L. B. and Resnick, D. P. (1992). "Assessing the thinking curriculum. New tools for educational reform", in B. Gifford, & M. O'Connor (Eds). *Changing assessment: Alternative views of Aptitude, Achievement and Instruction*, London, Kluwer Academic publishers.

Robson, C. (2007). *Real world research. A resource for social scientists and practitioners – researchers*. Second Edition. Oxford, UK: Blackwell Publishing.

Saldaña, J. (2013). *The coding manual for qualitative researchers*. Second Edition. Los Angeles: Sage.

Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education, 92*(3), 447–472. <http://dx.doi.org/10.1002/sce.20276>

Selghed, B. (2006). Betygen i skolan – kunskapsyn, bedömningsprinciper och lärarpraxis [Grades in the school – Knowledge views, assessment principles and teachers' practices]. Stockholm: Repro 8 AB Nacka.

Silverman, C. (2010). *Doing qualitative research, a practical handbook*. Third Edition. London: SAGE.

SKOLFS. (1992). 1992:4 Skolverketsförfattsamling [The Swedish National Agency for Education].

Skolverket (2012). Curriculum for the upper secondary school. Natural Science Programme, 2012-10-31 [The Swedish National Agency for Education]. Retrieved from www.skolverket.se

Skolverket (2013). Curriculum for the upper secondary school [The Swedish National Agency for Education, Stockholm]. Retrieved from www.skolverket.se

SOU (2008). Statens Offentliga Utredningar 2008:27. [The Swedish Government Official Government Report].

Stiggins, R. (2006). Assessment for learning: A key to motivation and achievement. Edge: The Latest Information for the Education.

Svensk författningssamling. (2010). SFS nr: 2010:800 Utbildningsdepartementet, 2010-06-23 [The Swedish Ministry of Education SFS nr: 2010:800].

Swedish Ministry of Education. (1994). Läroplaner för det obligatoriska skolväsendet och de frivilliga skolformerna [The Swedish Curriculum, 1994].

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education*, 13(3), 239–264.

Treagust, D. F., Jacobowitz, R., Gallagher, J. L., & Parker, J. (2001). Using assessment as a guide in teaching for understanding: A case study of a middle school science class learning about sound. *Science Education*, 85(2), 137–157. [http://dx.doi.org/10.1002/1098-237X\(200103\)85:2<137::AID-SCE30>3.0.CO;2-B](http://dx.doi.org/10.1002/1098-237X(200103)85:2<137::AID-SCE30>3.0.CO;2-B).

Wiliam, D. (2000). Formative assessment in mathematics. Part 3: The learner's role. *Equals: Mathematics and Special Educational Needs*, 6(1), 19–22.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).



Everyday classroom assessment practices in science classrooms in Sweden

María del Carmen Gómez · Anders Jakobsson

Received: 14 April 2014 / Accepted: 14 April 2014 / Published online: 4 June 2014
© Springer Science+Business Media Dordrecht 2014

Abstract The focus of this study is to examine to what extent and in what ways science teachers practice assessment during classroom interactions in everyday activities in an upper-secondary school in Sweden. We are science teachers working now with a larger research project on assessment in science education that seeks to examine teachers' assessment practices in the upper-secondary school. Framing questions include: are teachers performing an integrated assessment of students' skills as the national curriculum mandates? If so, what do the instructional discourses look like in those situations and what are students' experiences regarding their agency on learning and assessment? We emphasize the social, cultural and historic character of assessment and sustain a situated character of learning instead of the notion that learning is "stored inside the head". Teacher led lessons in three science classrooms were video-recorded and analyzed by combining ethnographic and discourse methods of analysis. Both methods are appropriate to the theoretical foundation of our approach on learning and can give some answers to questions about how individuals interact socially, how their experience is passed on to next generations through language and how language use may reveal cultural changes in the studied context. Making the study of action in a classroom the focal point of sociocultural analysis supports the examination of assessment processes and identification of the social roles in which teachers and students are immersed. Such an approach requires observations of how teachers act in authentic teaching situations when they interact with their students in classroom making possible to observe negotiation processes, agencies when both teachers and students are involved in every-day activities. Our study showed that teachers mostly ignored students' questions and that students solved their own problems by helping

Lead Editor: Alejandro J. Gallard M.

M. C. Gómez (✉)

School of Science Education, Malmö University, Valdemarsväg 3, 224 74 Lund, Sweden
e-mail: maricarmen7879@gmail.com

A. Jakobsson

School of Science Education, Malmö University, Hjälmarsväg 10, 205 06 Malmö, Sweden
e-mail: Anders.jakobsson@mah.se

each other. Teachers did not provide opportunities for students to discuss or argue scientific issues as the national science curriculum stipulates. We found that traditional assessment methods, such as tests, examinations and assignments were the most common methods used to assess and grade students' learning. Different aspects of knowledge stipulated in the national Swedish curriculum, such as lifelong learning, stimulation to students' creativity, curiosity as well as their wish to explore and convert new ideas into action, and find solutions to problems, were restricted by teachers' discourses. The observed teachers' learning and assessment practices constrain students' agency leading to students' silence consequently hindering students' development.

Abstrakt Fokus för denna studie är att undersöka i vilken omfattning och på vilket sätt lärare i naturvetenskapliga ämnen betygsätter eleverna under vardagliga aktiviteter och interaktioner i klassrummet på en gymnasieskola i Sverige. Vi är lärare i naturvetenskapliga ämnen på gymnasiet och arbetar nu i en större undersökning med syfte att ta reda på gymnasielärarnas bedömningspraktiker. Våra frågor inkluderar: utför lärarna en integrerad bedömning av elevernas kunskaper som de nationella läroplanen fastställer? Om så är fallet, hur ser diskurserna ut i dessa situationer och vilka är elevers erfarenheter kring sitt eget deltagande i lärande och bedömning? Vi betonar den sociala, kulturella och historiska karaktären av lärande och bedömning, och stödjer den situerade karaktären av lärande i stället för det lärande som finns "lagrat inne i huvudet". Lärarledda lektioner på tre naturvetenskapliga klasser videoinspelades och analyserades genom att kombinera etnografisk och diskursanalysmetod. Båda metoderna är lämpliga avseende den teoretiska grunden för vårt synsätt på lärande, och kan ge några svar på frågor om hur individer interagerar socialt, hur deras erfarenheter förs vidare till kommande generationer genom språket, och hur språkanvändningen kan avslöja kulturella förändringar i den aktuella kontexten. Genom att fokusera på handlingen använder vi ett sociokulturellt perspektiv, som möjliggör en undersökning av bedömningsprocesser och identifikation av de sociala rollerna som lärare och elever spelar under bedömningssituationer. För detta behöver vi observera hur lärarna agerar i en autentisk undervisningssituation, där de interagerar med sina elever i klassrummet, som gör det möjligt att observera förhandlingsprocesser, eget deltagande i lärande och bedömning, när både lärare och elever deltar i dagliga aktiviteter i klassrummet. Vi fann att lärarna oftast ignorerade elevernas egna frågor och att eleven löste sina egna problem, genom att hjälpa varandra. Lärarna gav inte eleverna möjligheter att diskutera eller argumentera vetenskapliga frågor, som naturvetenskapliga läroplanen stipulerar. Vi fann att lärarna använder traditionella bedömningsmetoder, såsom skriftliga prov, examinationer och läxförhör, och att dessa var de vanligaste metoderna för bedömning och betygsättning. Vi har i den här studien konstaterat att frågor om lärande, som anges i den svenska läroplanen; livslångt lärande, stimulans för elevernas kreativitet, nyfikenhet och deras önskan att utforska och omvandla nya idéer till handling, samt finna lösningar till problem, begränsades av lärarnas diskurser. Lärarnas lärande och bedömningsmetoder begränsar elevernas medverkan. Vi anser att detta leder till att eleverna tystnar, vilket i sin tur hindrar elevernas utveckling.

Keywords Assessment and learning · Science education · Classroom interactions · Language and power

Nyckelord Bedömning-lärande · Naturvetenskapen · Klassrumsinteraktioner · Makt och språk

In the last few years the process of grading-assessment in upper secondary schools has been a focus of public debate in Sweden. The center of attention in these debates is grade equivalence, students' low performance rates and the difference between students' increasing grades and decreasing scores on international large-scale studies. The underpinnings of discussions about grades in the media are economic and political especially as they converge with school accountability and competition between schools. Five years after the implementation of a major national education reform effort, Swedish newspapers were debating a report published by the Swedish National Agency for Education. The report showed that one out of four students in upper-secondary schools failed to pass one or more of the eight compulsory subjects (Sydsvenska Dagbladet 4 Juni 1997) [Southern Swedish daily 4 June 1997]. The eight compulsory subjects are English, History, Physical Education and Health, Mathematics, Science, Social Studies, Swedish and Swedish as a Second Language (Skolverket 2013) [The Swedish National Agency for Education 2013].

The columnist Helén Gustavsson wondered if it was the lack of knowledge among students or the assessment system that failed to reach the goals mandated in the education reform movement. One of these goals was "to ensure development of students' knowledge to prepare them for vocational diploma or studies in higher education; which means that the student has achieved a level of professional expertise providing good preparation for a professional life and for a life as adult and as a citizen responsible for their own lives." (Ministry of Education 1994, p. 27). She suggested that the problem was not to be found with the students, but that something had changed in the assessment system of 1994, which could explain students' low performance rates in the eight compulsory subjects listed above. She also suggested that in order to understand this issue one should bear in mind that the new assessment system in Sweden was introduced at the same time that major changes occurred in Labor market policy thus making a connection to neoliberalism, a subject we will discuss later. These changes and the implications for society have been summarized by Donald Broady, Mats B. Andersson, Mikael Börjesson, Jonas Gustafsson, Elisabeth Hultqvist, and Mikael Palme (2000) in a report entitled: *Skolan under 1990-talet-sociala förutsättningar och utbildningsstrategier*. [The school in the 1990s-social conditions and educational strategies.] According to the their report, the changes in the Swedish school system during the 1990's include goals and accountability focused around issues of decentralization, privatization and marketization. This direction involved changes in the whole educational system and reveals major ideological changes in Swedish society.

According to Magnus Dahlstedt (2007) changes in Swedish society over the past two decades are based on two central themes: decentralization and freedom of choice. These two primary themes have impacted the educational system, evolving "from collective to individual, from public to private, from outer to inner demands. During the same period the education system has shifted from top-down government to freedom of choice and individual responsibility, from conformity to diversity, from practices of direct control to practices of accounting, from sanction to self-discipline." (p. 64). Because of the aforementioned changes in Swedish educational policy the processes of control and management of schools has resulted in schools competing with each other—partially through privatization of schools—and the freedom for students to choose a school. In this system, students are expected to take more responsibility for their learning and more responsibility for education is placed on the family (Dahlstedt 2007). Student failure can be linked to low socioeconomic status (SES), gender, and ethnicity and other similar contextual factors in the educational system and society at large. Following a world-wide pattern, there is an increasing division between rich and poor, threats to environment and economic insecurity for low SES families. Both students and families are forced to assume the burden and

consequences of the limited school choices they are afforded and unfortunately blame themselves for failure. School choice for these students and families is not to be confused with those of parents who have the means to choose any school in the Swedish education system.

These two decades of reform have led to dramatic changes in Swedish education. Some of these changes include, growing ethnic segregation and class differentiation, both social phenomena enabled by the principle of freedom of choice making it possible for students from well-educated families that also have financial means to choose less socially vulnerable and immigrant dense schools. This student movement was not possible before the reform because the Swedish social system had mechanisms to counter this based on the principles of social welfare and harmony. “Since 1980, the educational system has increasingly come to function as a platform for individual development rather than as a redistributive tool” (Dahlstedt 2007, p. 63). The changes during these two decades have also led to more cultural and economic injustices. For example, according to Broady and colleagues (2000), by 1990 science programs had consolidated their status as the most important elite education programs in Swedish upper-secondary school, where only few students with lower SES succeed in entering.

According to Dahlstedt (2007), however, the differentiation of pupils in schools follows intrinsically the same pattern as before while the reforms during the last 20 years have reinforced the prevailing situation of inequity in the school system of Sweden. Consistent with the neoliberal agenda of making teachers more accountable, changes in the school system, including assessment, were required by the Swedish government as necessary for advocating for more government control than before the reforms. Dahlstedt (2007), argues that the need to make Sweden more competitive in the world and adapt to changes in the labor market were the key arguments from the government for the aforementioned changes in the education system. Test results, for example, can now be used as a selection instrument when students choose a school, as a control of a school’s goals, and as an argument for state authorization of opening new schools. In the neoliberal context, compulsory examination tools become redundant, the central test has been removed and grades have become an important foundation for education policy decisions. This change in emphasis influences teachers’ assessment practices as they must accommodate and transform learning and assessment according to the changes required by the Swedish government. The neoliberal ideology behind reform processes not only encourages reconsidering the purposes of assessment, but also leads to a shift of assessment procedures.

Neoliberalism

We shall refer to the term neoliberalism as a political project of governments to facilitate the re-structuring and re-scaling of social relations in different domains according to the requirements of an unrestricted global capitalism (Fairclough 2003). In order to sustain the structural continuity of the capitalist system, governments have engaged in radical restructuring, involving “both ‘re-structuring’ of relations between the economic, political and social domains (including the commodification and marketization of fields like education—it becomes subject to the economic logic of the market), and the ‘re-scaling’ of relations between the different levels of social life—the global, the regional (e. g. the European Union), the national and the local.” (Fairclough 2003, p. 4). According to Norman Fairclough, language is unescapably involved in these new processes of neoliberal

order because language is “an irreducible part of social life, dialectically interconnected with other elements of social life, so that social analysis and research always has to take account of language” (Fairclough 2003, p. 2). Consequently the neoliberal order permeates our society and schools and this process involves new ways of being and acting, and new ways for using language.

Understanding the influences of neoliberalism, many researchers define the purpose, the use, means and interpretation of assessment in a broader perspective, as we do also in this research. Specifically, we recognize that assessment is embedded in issues of culture, politics, ethnicity, SES and power. For example Patricia Broadfoot (1996) considers that assessment has emerged in society to help shape and organize education for all in response to major changes in society resulting from a neoliberal framing agenda.

According to Broadfoot (1996), assessment procedures mediate the relationship between education and society in contemporary education systems. Mass education and formal assessment appear to be inseparable in modern societies because the underlying rationale of all kinds of formal assessment is the neoliberal ideology involving issues of social control of education by the state. She pointed out various trends in the relationship between education and society including: proof of competence, control of competition and content and control of individuals and the educational system itself. Accordingly, assessment approaches have a crucial role to play in the Swedish system of education by attesting to competence, regulating competition and diminishing the individual’s frustration of those who have not succeeded in realising their objectives. There is evidence that these trends are also identified in many industrialized countries: “towards the certification of more industrially relevant competencies; towards a decline in formality in individual pupil assessment as it becomes more the responsibility of teachers and selection imperatives give way to more inclusive priorities; and, at the same time, a corresponding increase in other external forms of system control, such as national monitoring and centralized curricula.” (p. 63).

The Swedish public debate on grading intensified during the 1990s and continued in the 2000s, as the presence of “grade inflation” was a centerpiece in government reports (e. g. SOU 2004:29) [State Public Inquiry SOU 2004:29]. Christina Wikström (2005) investigated the mechanisms behind grade inflation and she concluded that grade inflation results in instability in the criterion-referenced grade system and in discrepancy of grades as selection instruments for school choice. She argued that increased grades could be explained by better achievement from students, however the increasing of grades cannot explain why incoming college students were continuously showing declining skills in mathematics. She suggested that the standards have been lowered either because teachers have problems with the current criterion-referenced grade scale or due to school competition between each other based on grades such that teachers were not able to endure the several forms of pressure for high grading. According to her, competition for higher grades is between teachers, schools and municipalities and the need to report higher grades to the Ministry of Education by reducing the proportion of failing students.

Wikström and Wikström (2005) point out that the decentralized structure of the school system limits central control mechanisms and that leaves the responsibility for grading students entirely to the schools and subsequently to individual teachers, who in turn defend their classroom assessments. They noted that there are no compulsory examination tools, such as standardised tests or external examiners, to judge the grade-setting process, and that grades are determined in the municipalities and conclude that this system leaves open the way for local grade inflation. Obviously, the immediate implication of grade inflation is the instability in the criterion-referenced grade system and the resulting discrepancy in selection for higher education and in the labour market, as the writers claim.

But assessment regulates competition. As Broadfoot (1996) argues: “Certification and the associated processes of selection has arguably long been the most commonly recognized function of educational assessment, since it involves students demonstrating their achievement in relation to the goals of the educational system.” (p. 32). A consequence of the divergence in grade instability of predetermined criteria of assessment and competition is also a divergence in every norm-referenced grade system. Students’ performances are evaluated in fair, objective and technically correct methods by competent and qualified teachers, “The subsequent ranking of candidates in comparison with their fellow competitors against predetermined criteria allows further and higher educational institutions and employers to select those whom they consider to have ‘performed’ the best” (p. 32).

Because assessment of students’ performances involves both competence and competition, education and proof of education through grading, are powerful instruments for controlling both the development of individuals and of society. We argue that individuals should be allowed to compete on an equal basis to show their competence. In addition if grade inflation occurs, a teacher’s assessment cannot be considered a neutral measure of a student’s merit.

Alli Klapp Lekholm (2010) claims that, “a fundamental problem with assessment in Sweden is the difficulty of attaining comparability of scores because it depends in the final analysis on teachers’ assessment practices and what they account for when they assess students’ learning” (p. 129). While we accept her affirmation that the grading system in Sweden has problems with grade calibration and that teachers assess students in different ways, we consider that these problems are merely technical issues of assessment that have risen as result of inherent contradictions caused by school reformation.

What Lekholm and other authors do not recognize is that teacher assessment is embedded within issues of culture, ethnicity, politics and power. Swedish teachers now are assessing students under a different kind of pressure, which is aligned with competition between schools and municipalities. A shortcoming in the prevalent research on assessment in Sweden is that it does not recognize the strong social and cultural character of assessment. For example, Eva Forsberg and Viveca Lindberg (2010) conducted a mapping of Swedish research on assessment, commissioned by The Swedish Council of Science Research, in which they asserted, “Our interest applies primarily to assessment as an object for scientific knowledge, not assessment as a social phenomenon, even if they are connected to each other. In scientific activity, it is the object assessment that is in focus for our activity” (p. 54) [*Vår intresse här gäller i första hand bedömning som ett objekt för vetenskaplig kunskap, inte bedömning som social fenomen, även om de är förbundna med varandra. I den vetenskapliga verksamheten är det objekten som är i fokus för vår verksamhet*].

What is it we are trying to understand?

This background brings us to the focus of the current study, which is to examine teachers’ assessment practices where they are interacting with students in everyday classroom activity. What is the significance of such interactions in classroom? In an assessment research context it is assumed that learning and assessment of the students’ performance is dependent on interaction in the classroom as the primary source of information (e.g. Broadfoot and Black 2004). However, this is a limited view because there are several kinds of assessment activities in a classroom. For example, sharing criteria of assessment with

students, collaborative assessment approaches among peers and peer- and self-assessment have been suggested to have potential effectiveness for learning (e.g., Black and Wiliam 2009). These activities are important for teachers to develop in their classrooms, together with their students, that is, making the interaction between a teacher and student an interesting pedagogical issue to be examined.

According to Caroline Gipps (1999) the processes of learning and assessment are deeply embedded and situated in cultural and historical contexts, but these situations are not normally taken into consideration within traditional Western assessment practices. This idea is in line with the findings of several researchers in the field of assessment. Paul Black and Dylan Wiliam (1998) noted, for example, that in studies carried out all over the world there exists a relatively strong empirical indication that the dominant picture of assessment processes in the science classroom is one where, above all, individual written tests are strongly entwined with the process of grading, and the way teachers assess students' learning determines the types of grades we give them. Wynne Harlen (2008) completed this picture by pointing to the use of everyday assessment methods, such as observation and the interpretation of science students' performance in the experimental parts of the instruction. A study by Martin Nystrand, Adam Gamoran, Roberth Kachur, and Catherine Prendergast (1997) demonstrated the significance of informal practices for assessing students' conceptual understandings, carried out by means of teachers' implicit or explicit evaluations of students' responses and questions in everyday interactions in the classroom. This interaction may be described as a sequential classroom organization, involving alternations of verbal and non-verbal behavior between the teacher and students, which often takes form in a classic initiation–student response–teacher evaluation (IRE) (e.g., Mehan 1979). According to Jay Lemke (1990), this pattern of interaction is culturally acquired, habitual, intuitive, and constitutes a tool with which to master the classroom and evaluate students' knowledge and performance.

In essence, we are interested in the pedagogical changes that the new curriculum requires, including changes about what knowledge is and how teachers assess knowledge. One question in our study is about how, and the extent to which, teachers use these kinds of classroom interactions as forms of assessments for grading students' performances. One obvious approach in relation to this focus is to conduct classroom observations and in-depth interviews with the teachers and students involved. By using discourse analysis in combination with an ethnographic approach we intended to describe teachers' assessment processes and practices over a period of time.

Thus, the focus on this study and the research questions can therefore be expressed as follows:

- To what extent and in what ways do science teachers practice assessment during classroom interactions?
- What do these assessment activities look like in a classroom in the context of the new Swedish curriculum of 1994 and the revised curriculum in force since 1 July 2011?
- Are teachers performing an integrated assessment of students' skills?
- What are students' experiences during the actual interactions at the lessons regarding their own agency on learning and assessment?
- Are teachers assessing students' achievement through interactions with them? If they are, then what do the instructional discourses look like in those situations?

Our study is one of several reporting the findings of a larger research project that aims to examine teachers' assessment practices in upper-secondary science classes in Sweden. The

larger research project is being performed in two stages. The first one was a study to investigate science teachers' assessment practices in the upper-secondary school in Sweden. In this stage, science teachers from different communities in southern Sweden were interviewed and asked about their assessment and grading practices (Gomez and Jakobsson 2013). The current study is a part of the second stage of the research project, which aims to closely examine teachers' assessment practices in interactional situations with their students. This article reports the results of the observations and in-depth interviews with the teachers involved in the current study.

Fairclough (2001) argued that the social condition of interactions in an institutional context is determined by conventions of discourse, but he also asserted that those conventions do not prevent individuals from being creative in the course of their everyday work. Our intention in the current study was to use his approach to discourse analysis to portray the dominant teacher discourses that have relevance to learning and assessment processes in classroom interactions.

Theories on context

Our research is grounded in an analysis of the language used for assessment in instruction by teachers and students during their everyday interactions in the classroom. We focused on the discourses teachers and students engage in as part of science learning and assessment activities. We drew primarily on concepts elaborated by Fairclough (2001) namely, discourse, practice, power, and social conventions to analyze the language used by teachers in assessing student learning. Specifically, we were interested in examining discourses focused on an individual's participation in shared intellectual activity.

Fairclough (2001) considered language a form of social practice that conveys short and long-term causal effects on both individuals and society as a whole; language is part of the society, he asserted, a socially conditioned process, which interacts dialectically with other non-linguistic elements of society. Conversations between people create communication through which thinking with others is possible. But language is also significant for understanding the relations between individuals within institutions inasmuch as relations between individuals belonging to different social categories involve power. Power, in turn, precedes speech, because what individuals can say is situated within existing institutions whose norms and conventions determine what can and cannot be said (Fairclough 2001).

The terms *discourse* and *practice* are considered to be ambiguous by Fairclough (2001) because both can refer to either what people are doing in a certain situation, or what people usually do given a particular circumstance. Thus, both terms can refer to action or to convention. The ambiguity of these terms is also appropriate to our study, which placed an emphasis upon the social character of discourse and of practice: "the individual instance always implies social conventions—any discourse or practice implies conventional types of discourse or practice" (Fairclough 2001, p. 23). However, this ambiguity also suggests the social prerequisites for individual agency: "the individual is able to act only in so far as there are social conventions to act within. An implication is that people are enabled through being constrained: they are able to act on condition that they act within the constraints of types of practice—or of discourse" (Fairclough 2001, p. 23). Discourse and practice are constrained by interdependent systems called "orders": "orders of discourse" and "social orders" (Fairclough 2001, p. 24). In this study we were concerned in understanding how teachers act within those orders of discourse during encounters in the

classroom with their students, and the way those practices contributed to the assessment of students' learning.

The school is a social arena where discourse happens, and every individual involved in school activities has a recognized social role, or *subject position*. The different kinds of discourses are established depending on the subject positions of teachers and students, and it is only by living within these positions that individuals become a teacher or a student. Thus, "occupying a subject position is essentially a matter of doing (or not doing) certain things, in line with the discursial rights and obligations of teachers and pupils—what each is allowed and required to say, and not allowed or required to say, within particular discourse type." (Fairclough 2001, p. 31). However, this kind of social determinism does not prevent individuals from being able to "de-structure" and "restructure" orders of discourse. The subject has the paradoxical characteristics of "...being socially determined, and yet capable of individual creativity; obliged to act discursively in preconstituted subject positions, yet capable of creatively transforming discourse conventions" (Fairclough 2001, p. 140).

What is the phenomenon we claim to assess when assessing students' outcomes? A sociocultural view of learning and assessment

Much research on assessment seems based on assumptions emerging from a quantitative view of learning and knowing and a predictive, deterministic psychometric framework (see Delandshere 2002). According to Ginette Delandshere (2002) most of the educational assessment specialists are still working with century-old conceptions and behaviorist approaches and in spite of the call for changes, educational assessment remains unchanged. She claims that with a behavioristic approach learning entails an accumulation of knowledge and learning comes from teaching, which is arranged to transmit pre-determined knowledge in a fragmented, sequential and hierarchical manner. The task of assessment is then to observe whether individuals can repeat this knowledge when they are stimulated to do so. We call into question this approach and propose a quite different learning and assessment method grounded on a socio-cultural perspective of learning and assessment, that includes historical, socio-cultural, and activity theories of learning. Thus, we follow Vygotsky's key feature of sociocultural view of learning which is that individuals learn through dialogue with others and that those ideas in dialogue with others appear first in an external social plan that then become internalized by the individual (Vygotsky 1978).

The basis of the sociocultural theory of human thinking and development was articulated by the early twentieth century thinker Lev Semyonovich Vygotsky. The contribution of Vygotsky and his colleagues to understanding human thinking and development is crucial, as it helps bridge the Cartesian dualistic line of reasoning, which separated body and mind (e.g. Wertsch 1991). A fundamental characteristic of sociocultural theory is its focus on human action (Wertsch 1991), and on humans using "mediational means" such as tools and language. From this perspective, individuals are regarded as being active developing agents and their development takes place through the use of tools such as language and their autonomy as learners. The interaction between individuals with mediational means may result in developmental changes within the agent. Language is an instrument for learning and modes of thinking and patterns of language use are social processes. However, focusing in language as a singular instrument for learning is not enough to understand the development of students' identities as learner. In education,

knowledge and pedagogy are closely interlinked to identity. None of these terms can be understood about without the others (Bernstein 1996).

According to a socially situated view of learning (e.g. Lave and Wenger 1991) learning is not only a cognitive activity, but also entails developing a social identity related to practice and discourse inside a community as well as across communities. Social identity can be theorized as related to groups to which individuals belong to and, as Fairclough (2001) argued, their roles within these groups. Accordingly, individuals will learn and demonstrate skills, knowledge and language use if they consider that they are part of the community and conversely they will fail to acquire knowledge and demonstrate knowledge if they not feel involved in the community.

Some theoretical aspects of these sociocultural theories on learning have inspired the current study in particular the social character of language. Individuals use language as mediational means, as means of social interaction, and/or a means of statements and understanding. Learning is not conceivable without social experiences with others and its dependence upon cultural resources. Following sociocultural theory of learning in our study goes beyond “within the head” models of learning because these ignore cultural contextual factors on learning and thereby assessment. A classic pattern of interactions in a classroom is that the teacher’s attention is focused on the “correctness” of student’s response.

The IRE pattern mentioned previously has been criticized because teachers who use the IRE pattern as the only or dominant procedure in the classroom have a propensity to be teacher-centered, and this approach fails to provide students with opportunities to voice their own ideas or comment on those of others. Gordon Wells (1993) argued that in this pattern of interactions, the teacher, by virtue of his/her status as primary actor or knower, both initiates and finishes the interaction. William Carlsen (1997) has also examined the tension between scientific argumentation and the “authority of knowledge” versus the “authority of position” (p. 15) and found that the teachers’ questions can play a role in discouraging students’ discourse. Wolff-Michael Roth (1996) argued that the evaluative sequence of IRE is seldom based on students’ genuine questions.

Kenneth Tobin and James Gallagher (1987) closely studied and documented what teachers and students actually do in science classes and found that the level of cognitive demand placed on students in classes and laboratories was likely to be relatively low; they suggested that in this dynamic of interaction, the possibilities of student engagement were limited by teachers’ management of the discourse. Several other observational studies (e.g., Carlsen 1991) have indicated that when teachers ask more authentic questions and when the question is less likely to take the form of an assessment, the students seem to produce longer and more complex answers and are more likely to contribute their ideas or opinions. Eduardo Mortimer and Philip Scott (2003) have come to the same conclusion. They worked with the IRE pattern in order to explain and develop a dialogical framework for analyzing science classroom communication and suggested that the emphasis ought to be less on the questions teachers pose and more on the way teachers react to students’ responses and genuine questions.

Classroom interaction research also suggests that teachers’ pedagogical practices may influence students’ participation in classroom discourse, or lack thereof (e.g., Mack 2012); that peers may also be involved in the dynamics of classroom interactions, for example, by controlling the discourse (e.g., Thornberg 2010); and that the teacher’s role in encouraging dialogic interaction in classrooms may improve students’ inclusion in science discourse (e.g., Martin and Hand 2009).

Our research design derives from real life in schools

The primary source of data was transcriptions of audio- and video-recorded lessons of three regular teachers teaching chemistry, biology, and physics in their respective classrooms; recorded statements from in-depth interviews with the teachers involved in the observations; and personal field notes. We followed the ritual chains of interactions and transactions that are deployed in everyday classroom activity in upper-secondary science classrooms and analyzed the collected data by carefully combining our ethnographic work with discourse analysis.

This procedure followed an ethnographic approach rather than a full-scale ethnography. The approach is termed “observer-as-participant” by Colin Robson (2007). In this approach the researcher “is someone who takes no part in the activity, but whose status as researcher is known to the participants. Such a state is aspired to by many researchers using systematic observation” (p. 319). The participation of the observer involves the observation of events, interactions, and behaviors, and the production of field notes. All observations were audio-recorded and video-recorded as suggested by Martin Hammersley and Paul Atkinson (2007). Specifically, they describe technicalities about the correct placement of the voice recorder, the cameras, light conditions, observers’ position, and so on.

The three classes were two classes from the first year courses in biology and chemistry; each class comprised about 30 students all aged 15–16 years; and a third class engaged in the 3rd year of a physics course that comprised 30 students aged 17–18 years. We completed a total of 20 observations as a result of which 50 hours of activities were recorded continuously in the three classes in sequential lessons over a period of four weeks; these lessons covered most of the activities students and teachers were engaged in, that is, both theoretical lessons and experimental work. Separately and in parallel with teachers lecturing, we focused on recording specific instances of student–student conversations during the lessons and laboratory activities.

At the end of the observations the teachers and students involved were asked to participate in in-depth interviews. The aim of the interviews was to find consistency in what had been observed and noted during the lessons in order to reach an understanding of the teachers’ actions in specific situations. We asked teachers about their beliefs about special events observed during their interactions with students and about the probable causes of an outcome.

We wanted to know what teachers and students were doing, and why through analysis of the videos, recordings and interviews

A review of the video and audio-taped records of all observed activities led by teachers in the classrooms and of the interviews with teachers, in two cycles of coding using the Verbal Coding Exchange System (Saldaña 2013) was conducted. Situations where we interpreted the interaction as an assessment event involving students’ learning were analysed. The first cycle coding method proceeded with an exact transcript of the vocal exchange, including non-verbal prompts and pauses between the speakers. The data was analyzed through the lenses of situated action and learning that concern the everyday activity of persons acting in [a] setting (Lave and Wenger 1991).

We chose and wrote out by hand those patterns of communication that gave us information about how teachers act when teaching science, paying careful attention to those acts. We drew from the transcripts a variety of forms of verbal exchange to categorize the

unit(s) as a pattern of “skilled conversations” (Saldaña 2013, p. 137) representing information exchanges among individuals, to include debates, inquiry, contradictions between teachers and students, and acts of negotiation. At this level we reflected on our inquiry and our aim to find patterns of interaction that facilitated or prevented students’ participation in both the learning and assessment processes. By analyzing their teaching and assessing the taught material, we also tried to find patterns of communication that showed how teachers assisted students’ learning. Student’s agency and engagement in learning and assessment were also analyzed carefully, including how both teachers and students’ exerted their roles in the subject; for example, how the students engaged or did not engage in science discourse with teachers and peers, and how teacher answered and assessed questions from students. The analysis focused specifically on how students’ ideas were evaluated and shared by teachers and peers, and on the way in which teachers provided opportunities for students to discuss, share, and negotiate students’ ideas and interest in the subject matter.

In a second level of the verbal coding exchange, we examined the teacher and student meanings of key instances by examining speech characteristics, such as the tone of voice of teachers and students, and non-verbal communication. Before the categorization proceeded, we reorganized and reanalyzed the data that was coded through the first cycle method and we then proceeded to the second cycle coding method. A second cycle coding helped us to develop a sense of the categorical organization of the data from the grouping that had resulted from the first cycle of verbal coding. In the second cycle we rationalized the data and developed a smaller list of broader categories, themes, and assertions. Then the categorization proceeded.

We noted during our ethnographic work that discourses concerning traditional assessment approaches, such as those about written tests and examinations, were ubiquitous in everyday classroom activities. Those discourses were not the focus of our analysis of the collected data. We were interested in examining the process of assessment involving activities that may be difficult to evaluate by traditional paper-and-pencil examinations, such as, aptitude for critical thinking, argumentation, knowledge acquired outside the school, and analytical skills. All of the aforementioned skills are mandated by the national curriculum and as such are to be regarded by teachers when assessing students’ performance. Unfortunately teachers place strong emphasis on paper pencil tests as evidenced by our study. However, we need to further research the reasons for this seemingly wholesale ignoring of mandated education policy.

As part of our ethnographic observations, we also took field notes with the explicit purpose of describing in words what was unsaid and unspoken. The field notes were neither a translation of experienced activity nor a final text, but a personal diary about what happened during every day of the observation (Goodall 2000). We produced an account of discursive activities among students; that is, which students were active participants in discursive activities and how often, which students did not participate in discourse, and which students always answered teachers’ questions correctly and often. The field note work was also a valuable tool for describing the classroom context in everyday observations.

Exclusionary practice and developmental conversations: science for some students

This section begins with an example of exclusionary practice observed in compulsory school. We describe in this narrative how a students’ failure may be linked with the way they are portrayed at school through for instance pedagogical discourses. The narrative has the intention to explain one of the causes why it is necessary to investigate the encounters between students and teachers in assessment situations.

The backdrop for this narrative is a mandate of the Swedish Education Department since 1994 called “developmental conversations”. It means that at least one time in each term teachers must have a conversation with students and their parents about their scholastic development or progress in most of the Swedish school levels. The specific purpose of such conversations is to help students improve their learning and their lives. This governmental decision has become legitimized in several governmental documents (e.g., SOU 2000:1; Skolverket 2013) [The Swedish Government Official Report, 2000:1; The Swedish National Agency for Education 2013] and is embedded in the global educational reform changes in Sweden. In such documents, developmental conversations are emphasized as a method for strengthening citizen participation and influence in democratic processes and to get students and their families engaged and held accountable in the educational process. Moreover, “Students are forced to take greater responsibility for their own education and to bear the burden of the consequences for the choices they make—or they do not make. Parents and children who do not live up to these expectations and demands tend to be portrayed as problematic or deviant. The implication being that they have no one else to blame but themselves” (Dahlstedt 2007, p. 64).

One of the authors, Maria was invited to participate in one of those developmental conversation meetings. Although she was not directly involved in this student’s science learning, she was asked to represent the sciences and mathematics teacher’s team. Maria has been a mathematics and science teacher for 30 years. In all of her years of teaching she has never before been astonished and affected when witnessing a school meeting. Specifically, one of the most subtle forms of exclusionary practices took place in a developmental conversations meeting in which she attended and left her shocked because she could not believe that this could happen in Sweden.

Maria had participated in this kind of conversation as a parent, and thought that developmental conversations were a good thing. However, this time the author was sitting not as a parent but as a substitute teacher who was there for 2 weeks, together with a student called Peter, which is a pseudonym. Peter, a fluent Swedish speaker in the upper-compulsory school or the 9th grade, is from Iraq and had been resident in Sweden for 12 years. His native language is Arabic. This is an urban school with a blend of students regarding social class and ethnicity. Others attending the meeting were Peter’s parents, the head-teacher, who is a supervising teacher with complete control of the class, and a social consultant that evaluates Peter’s schooling and his social development.

Maria in a discussion with the co-author supposed that she was chosen by the head-teacher to be present in the meeting because of her position as a substitute teacher, without any power to make decisions. A tenured teacher, a teacher that is normally in charge of the science lesson, may possibly have the option of expressing her opinion about Peters’ situation, which should be discussed in the meeting, but in this meeting, as planned by the head-teacher no further input or opinions were sought from anyone in attendance because the outcome of the meeting was decided in advance.

Peter was sitting with his head down looking at the floor and the head-teacher began the meeting by reading in her papers a description about how Peter had failed the last physics test, and how he also failed the last term in a chemistry course and how horrible it would be if he also failed in biology. She told him that what he needed were good grades in mathematics, Swedish and English and that he did not need to study the sciences. She asked him if it would not be better for him to focus on those subjects. Peter responded by saying yes without taking his eyes off the floor. The head-teacher asked his parents to confirm her recommendation saying that he did not need sciences anyway and it would be good for him if we decide at this time that he dropped out of the science classes.

Maria reflected on the head-teacher's influence on a student's learning and life chances. The head-teacher used her power to construct ideas about how bad Peter was in science, and she likely covertly construed consensus on Peters' impossibilities to learn sciences and argued about his need for grades in only three subjects, Mathematics, English and Swedish. In the meeting nobody could argue against the head-teacher, nobody could say "no" to the head-teacher. She was correct in the eyes of the others in the meeting!

The parents agreed also and asked Peter, would you like to learn only math, Swedish and English? Yes, he said, with a resigned expression in his eyes. Peter felt perhaps that he was so bad in sciences that he should not contradict the adults. After this meeting, which was arranged to "help a student in his development", Peter was removed from science classes and placed together with other students being excluded from science classes in similar circumstances. "A", a teacher that knew of Peter, was angry because she argued that what the head-teacher did was not right and was not fair. The decision was unfair because it was a discriminatory act against Peter that resulted in excluding him from any future science learning.

This institutional example of decision making that leads to exclusionary practice raises the question about how Peter's science teachers used his way of knowing for learning science and whether Peter had been provided opportunities to recount his own everyday scientific experiences or knowledge ability outside science class. The idea behind tracking pupils in the compulsory school is to increase the proportion of students passing grades. This practice is in line with school policy to make teachers and schools more accountable on the basis of the proportion of passed students. The spirit of developmental conversations, which was proposed to provide a sense of scholastic direction, was ignored and instead used to maintain a *de facto* status quo as the new Swedish accountability educational system demands.

Alejandro J. Gallard M., and René Antrop-González (2013) differentiate between *de jure* and *de facto* aspects of education policy and argue that *de jure* educational policies, which are often legal guidelines that mandate minimum levels of compliance, unfortunately become translated to mean the normative way to implement educational practice: "...at the heart of the underpinning of *de facto* exclusionary practices is a total disregard for any practice that will enable the enfranchisement of people by going beyond the *de jure* aspects." (p. 995). The spirit of developmental conversations is to help students, as they progress through their academic career, to make informed decisions about education and their future. What has become *de facto* though is to use developmental conversations as a vehicle for exclusion of students with minimum if any student participation in the process. This echoes Gallard M and Gonzalez's belief that:

...*de jure* pedagogical actions that become *de facto* can and do serve as smoke screens for at the very least exclusionary practices. Specifically, what is viewed as doing the right thing as an effort to provide access for all, in reality does little more than meet the *de jure* aspects of teaching and learning. Indeed, this made teaching easier, but it does not address the spirit of the national curriculum, which is to meet the learning needs of the students. This is because the spirit is lost and replaced by *de jure* pedagogical actions, which are the most minimal of acts

(Gallard M., and Antrop-González 2013, p. 994). In this study developmental conversations became a normative implementation of an educational mandate that does not follow the spirit of the same. Thus a *de facto* status quo in school ensures that the law is followed by the teacher-head, even if she does not follow the spirit of developmental conversations; and secondly she ensures the implementation of the new Swedish school accountability educational reforms that specifically demand that all students will approve at least three subjects to continue studies at the upper-secondary school.

One of the inferences one can draw from this history is that when science students come to the Swedish upper-secondary school they are already differentiated from other students like Peter. In the upper-secondary school, this exclusionary practice endures and science is portrayed as difficult and intended for only students having special talents. For example, in an ethnographic study of classrooms interactions in the upper secondary school in Sweden, Dennis Beach (1999) observed different demands on students in economic and business programs than those in the science program. Comparing mathematics tasks of the science program to the other programs he observed that the complexity of tasks were different. He encountered differences even on working speed, attendance at school, punctuality, the way teachers see students and their own teaching, and recruitment to qualifications based education, which Beach noted, “privilege to already privileged students by the system” (p. 354). In science programs at the upper secondary school, according to Beach, it is not possible to work less if students are to stay in the program. Science programs have a mechanism, such as higher tempo, and the consciousness that students will demand more difficult mathematic drills in the future, preventing non-productivity.

The description provided in the preceding paragraph is an ambitious pedagogical strategy that results in students who have difficulty meeting the productivity demands dropping out the science and mathematics track and consequently being evaluated out. The process is not the same in the economic and business programs. According to Beach, these programs have their own mechanisms, including low-productivity as a norm. These findings suggest that science in the upper-secondary school is portrayed as demanding, intended only for high achieving students, and less available for all students. It could also mean that those who study economics or business are destined to be clerks or some other low status job, in summary, reflecting how the social reproduction in upper-secondary school is expressed as before the reform.

Eva Nyström (2007) has also reported inequality at the upper-secondary school in Sweden. Her research has made visible how these processes about teaching and learning science are constructed in talk or discourse situations. She draws on a range of post-structuralist methods that reject hegemonic and essentialist ways of seeing sciences, arguing that factors such as social categories, gender, ethnicity and sexuality are constructed during dialogues when teachers and students talk science and practice the same. Alejandro Gallard, and colleagues, (2013) refer to these influencing factors as contextual mitigating factors. In our study, we use the term CMFs to indicate that “there is no past, present or future action, existence, or thought that has not been mitigated upon by influencing factors that are part of all contexts. We do not argue they are defined equally by all contexts. Rather they are unique and defined by a particular context in which their influences are experienced.” (p. 4)

In addition to the view of science described by Beach (1999) at the upper level of schooling, pupils at lower grades also experience processes of inequality and exclusion. In other words, the problem of the misuse of assessments begins in the compulsory lower level and not the upper level. Learning is focused on the three subjects mandated for admission to upper secondary school: Mathematics, Swedish and English (Swedish national school curriculum, Ministry of Education 1994). This situation has been of some concern to the Swedish government, which formed a commission to investigate this matter (Skolverket 1998) [The Swedish National Agency for Education 1998]. The lowering of expectations for pupils has resulted in an increasing portion of pupils who fail to achieve one or more of the three focused mandatory subjects for admission to upper secondary school. It may well be that such students respond to lower expectations by identifying themselves as unfit for studies and consequently, they give up studying. The negative

outcome from lowered learning demands on pupils suggests that the narrowed focus of teaching is not for pedagogical reasons, but more a response to the new Swedish accountability system within and across schools.

In the new, revised curriculum for the upper secondary school, depending on the program that is chosen, grades from eight or twelve subjects such as mathematics, Swedish, science, English, mathematics must be at a passing level in order for the pupil to be admitted to upper secondary school (Skolverket 2013). [The Swedish National Agency for Education 2013] The education programs now offered are divided into two main groups, one preparing for further studies, such as science programs, that must have passing grades from twelve subjects. The other prepares students to be auto mechanics, electricians, auxiliary nurses or dental assistants requiring passing grades from eight subjects. An example of some of the subjects, which students must pass whether 12 or 8, are: Swedish, biology, chemistry, English, geography, history, mathematics, physics, and social science. The reason for the difference between the two programs, in terms of the number of subjects that a student must pass, is determined by the Ministry of Education. However, we believe that this difference in expectations is more about weeding out students under the guise of greater productivity demand. For example, those who are placed in health and care programs by their schools upon graduation immediately go to work and have little chance of studying at a university.

The revisions to the curriculum of 1994 may be interpreted as a formalization of exclusionary practices already developed within the upper secondary school, where different groups of pupils have access to different knowledge and standard of education. For example, a new regulation for the upper-compulsory primary school, called Regulation 2011:355, was announced by the Government, through the Swedish Education Ministry on the 2011-03-31 in which students were to be recruited into “spetsprograms” which are elite programs with specialized courses. This new regulation may be interpreted as reinforcing exclusionary practices by separating students into different classes in accordance with their academic ability especially in the sciences. Portraying science as an elite subject may influence a teacher’s assessment practices in the upper secondary school, and probably even influence students’ responses. However, the research on our assertion is very limited.

What we found in our ethnographic study

The results of the current research study may be summarized according to two major themes as follows:

- *Assessment practices by teachers in everyday activities in science classrooms:* Several features of the dominant discourses relevant to the assessment and grading of students are presented in this section of this article.
- *What the teachers said about their assessment practices in everyday classroom activities:* This section presents the teachers’ own statements about how they assessed students, and their conceptions of the way they assessed student learning and outcomes.

An analysis of the material obtained from the research study is presented next, enlarging on each of the two major themes summarized above. The actual situations in the science classrooms are set out based on the transcripts from our ethnographic observations and interviews with the teachers. The first major theme is related to the research questions because it explains to what extent and in what ways science teachers practice assessment during classroom interactions and how these assessment activities look like in a classroom in the context of the new Swedish curriculum of 1994 and the revised curriculum in force

since 1 July 2011. The second theme represents what the teachers expressed about their own assessment practices. This second theme is related to the requirement of the new curriculum, which is to assess all student knowledge, including those acquired outside the school context. These themes are related to the Swedish national curriculum of 2011.

Assessment practices by teachers in everyday activities in science classrooms

The transcripts A, B, C and D are related to the research questions in many ways. Transcript A shows to what extent and in what ways science teachers practice assessment during classroom interactions and how the instructional discourses look like in those situations. This is related to the following research question: To what extent and in what ways do science teachers practice assessment during classroom interactions? Transcript B shows what the assessment activities look like in a classroom in the context of the new Swedish curriculum of 1994 and the revised national curriculum of 1 July 2011. This transcript is related to the following research question: What do these assessment activities look like in a classroom in the context of the new Swedish curriculum of 1994 and the revised curriculum in force since 1 July 2011? Transcript C is an example of student's actual experiences during the lessons. This transcript is related to the following research question: What are students' experiences during the actual interactions at the lessons regarding their own agency on learning and assessment? Transcript D shows what the instructional discourse looks like. This transcript is related to the following research question: Are teachers assessing students' achievement through interactions with them? If they are, then, what does the instructional discourses look like in those situations?

Transcript A: The following transcript is from a lesson where the chemistry teacher, Maria, lectures on the electronic structure and characteristics of elements in the periodic system of elements.

Turn	Time (min)	Main dialogue Speaker	Parallel dialogue Speaker
1	6.32	Teacher: Maria ...and let's see... we know that here we should find in total eight electrons... and how should these electrons be distributed? (...) Helen! [...och vi får se ...vi vet att här ska vi hitta åtta elektroner ... och hur är de placerade? (...) Helen!]	
2	6.38		Charles Two on the first and six. [Två i den första och sex
3	6.40	Helen Hum... two on the K-shell and six on the L-shell. [Hum... två i K-skal och sex i L-skal]	
4	6.45	Teacher: Maria Yes...[ja]	Robin What? [Vad?]
5	6.47	Teacher: Maria ...and so, we go on... Yes...[...och vi fortsätter... ja ...]	Robin What did you say? [Vad sade du?]
6	6.49		Charles Two on the K-shell ... [Två i K-skal ...] (Sept 2009)

In this event, the teacher starts the dialogue by initiating (I) the conversation, asking “how should these electrons be distributed?” (Turn 1). Three of the students raise their hands and the teacher directs to Helen to respond (R): “hum...two on the K-shell and six in the L-shell” (Turn 3). Soon the teacher evaluates (E) the answer: “Yes...” (Turn 4). This structure of talk is similar to the IRE-pattern of interaction (Mehan 1979). The interaction shows one way science teachers practice assessment during classroom interactions. The extent the teacher assesses students is given by the extent to which the IRE pattern is used by the teacher because during the interaction the teacher is always evaluating students’ statements. In every interaction the students are held accountable. At this moment, however, we do not know how the teacher uses those assessment activities to grade students.

We do not consider that teachers are performing an integrated assessment of students’ skills when learning and assessing solely through this kind of interaction. In accordance with the stipulations on the new national curriculum in force since 1994, and the revised national curriculum, in force since 1 July 2011, when teachers award grades they should take into account the knowledge students have acquired outside the actual teaching situation and on the basis of the national knowledge demands for each course make a comprehensive assessment of each student’s knowledge. The questions teachers pose to students during lecturing are short, and are concentrated on the correctness of the answer. Students are not given the opportunity for analysis or reflection during those interactions thus making it difficult for teachers to perform an integrated assessment of students’ skills. In order to perform a comprehensive and multifaceted approach needed to assess students in accordance with the new national curriculum teachers are required to use different skills to capture evidence of learning from students. This task is more complex requiring a more multifaceted assessment strategy than a paper-and-pencil test. In our study beside the IRE pattern of learning and assessment, we have not found any other alternative method of assessment being used than paper-and-pencil tests.

During the dialogue sequence captured in Transcript A, other dialogues progress in parallel and silently between students. These dialogues are important pedagogical tools for students because they use them in collaboration with their peers to try and make sense of the subject and in so doing invite further comments by other students seeking to understand the idea of electron distribution in the atom. The dialogues between students also reveal that during IRE interactions students must in some way solve their own problems both individually and collaboratively. The dialogues may be also interpreted as showing the difficulty students experience asking teachers about science concepts and ideas they do need to share with to each other “What did you say?” (Turn 5).

Transcript B: What do these assessment activities look like in a classroom in the context of the new Swedish curriculum of 1994 and the revised national curriculum of 1 July 2011?

The teacher, Maria, is explaining at that moment the distribution of electrons in the transition elements. It is the continuation of the lesson shown in Transcript A on the periodic system of the elements. The main characteristic of this dialogue is that the teacher remains silent when students want to discuss the transition elements.

			Main dialogue	Parallel dialogue	
7	13.54	Louise	But how can palladium have zero valence electrons? Then it can't react... as you see...[Men hur kan palladium ha noll valens elektroner? Då kan den inte reagera... som du ser ...]		
8	14.00	Teacher: Maria	Palladium... [paladium]		
9	14.05	Teacher: Maria	[silence]	Charles	Do you understand? [Förstår du?]
10	14.08	Teacher: Maria	[silence]	Jon	I understand everything but palladium...I do not...[jag förstår allt men inte palladium.. jag förstår inte]
11	14.11	Teacher: Maria	If you ask me about the chemical characteristics of palladium... I do not know... eehh...but it seems that the outer shell is all occupied, isn't? [Om du frågar mig om de kemiska egenskaper av paladium... Jag vet inte ...men det ser ut att ha sin yttre skal hel fylld, eller hur?]	Charles	I don't get it either. [Jag förstår inte heller]
12	14.14	Louise	Pd... number 46, the book says that it has zero valence electrons ...how can it react when it has zero valence electrons? Our book says... 2, 8, 18, 18, 0 [Paladium är nummer 46 i tabellen, i boken står att den har 0 valens elektroner ... hur kan den reagera om den har noll valens elektroner?]	Jon	Neither do I. [Inte jag heller]
13	14.20	Teacher: Maria	Yes, exactly, and this has to do with the electron, which is added in each step to the orbital of lowest energy... therefore the new electron is placed in the inner [sub] shell, and therefore there are not electrons in the outer shell. [Ja just det, och det har att göra med elektronen som adderas i varje steg till skalen av lägsta energi ... därför den nya elektron placeras i de innersta skal, och därför finns inte elektroner i den yttersta skal].		
14	14.25			Charles	Say something! [Säg någonting!]
15	14.27	Louise	...then it could not react with ...consequently... [... då kan den inte reagera med ... följaktligen]	Jon	hum?
16	14.29	Teacher: Maria	[silence]	Charles	Say something... Say something! Säg någonting!]
17	14.39	Teacher: Maria	[silence]		[silence] (Oct 2009)

Student, Louise, breaks the IRE pattern of interactions by asking the teacher a question. Maria, the teacher, tries to answer Louise but is unsuccessful. Louise insists, “But how can palladium react when it has zero valence electrons” (Turn 7). The teacher formulates a new answer from another perspective that requires a higher cognitive level but Louise does not master it: “[it] has to do with the electron, which is added in each step to the orbital of lowest energy...” (Turn 13). This short, instantaneous discussion is finished by Louise’s hesitation: “...then it could not react with ...consequently...” (Turn 15). The teacher does not answer the question, she became silent and the dialogue ends at this moment. The students’ limited opportunities to talk or discuss concepts or epistemologies of science are revealed in this transcript.

The student–teacher dialogue could have been developed in a discussion about other approaches to explain the electronic structure, but as the transcript shows, the teacher closes the possibility for Louise to ask further questions by remaining silent (Turns 16–17). In this kind of interaction, students have limited opportunities to obtain answers to their own questions. The teacher’s talk dominates the episode and the rest of the students are silent. It may be understood as an interactive/authoritative pattern of the communicative approach (Mortimer and Scott 2003).

The teacher’s approach to learning and assessment of sciences as demonstrated in this excerpt, and characterized by the teacher limiting opportunities for students to talk or discuss various concepts of science, is in contradiction to both the new national curriculum of 1994 and the revised national curriculum of July 2011. In turns 16 and 17 of Transcript B, the teacher remains silent and in this way silences the students also. By this act the teacher does not promote students’ opportunities to influence their learning. In addition, when the teacher silences students they are denied the possibility to either ask questions or participate in their own development, undermining their agency. An individuals’ agency is important because in the *free choice* paradigm, school is supposed to be a place with opportunities for individuals to pursue careers and therefore the students strive to get good grades by asking the teacher questions that may help them in their development and lead to better opportunities in future life as individuals.

As Transcript B reveals, Charles and Jon are in the same situation as Louise. Neither Charles nor Jon understands the issue, but they do not say anything to the teacher or the class. Their attempts to understand the issue of transition elements seem fruitless and their opportunities to further discuss ideas are constrained by what they seem to perceive as their lack of understanding and their lack of access to the teacher. Charles does not give up and insists his peers to talk and to say something: “Say something... Say something!” (Turn 16). He does not dare ask question to the teacher possibly because he feels the environment does not allow him to talk and he stays silent.

Transcript C: This transcript shows what the teachers’ instructional discourse looks like during lectures in a specific learning situation. The teacher has recently ended the dialog with the class by remaining silent for some minutes. By remaining silent the teacher appeals to her authority and her position as a knowledge expert: the development of a non-interactive authoritative approach to science instruction and the monologic character of the interactions.

Main dialogue			
18	16.58	Teacher: Maria	I feel so!!!... I shall not enter into discussions about the transition elements because they do not behave as they should. And it has to do with Bohr's atom... it is not 100 %. (...). And because those transition elements are very troublesome, we leave them! [Jag känner såhär... jag ska inte komma in med diskussioner om övergångs element för att de bete sig inte som det ska. Och den har att göra med att Bohr's atom... inte är 100 %. (...). Och de ämnena är mycket besvärliga, vi lämnar de!]
19	17.55	Louise	But... [Men...]
20	18.00	Teacher: Maria	We leave the transition elements and focus on group III A (...) all the atoms strive towards getting 8 electrons in the outer shell... [Vi lämnar övergångs element och fokuserar på grupp III A (...) alla atomer strävar efter 8 elektroner i den yttersta skal...]. (Oct 2009)

The teacher eventually interrupts her own silence, continues the lesson and finishes the communication about transition elements. The teacher finishes the discussion by appealing to her authority: “they do not behave as they should. And it has to do with Bohr's atom... it is not 100 %” (Turn 18). The teacher's instructional discourse emphasizes her status as a primary actor or knower: “And because those transition elements are very troublesome, we leave them!” (Turn 18). Students' agency in this situation is variable, some of them became silent, others were talking to each other, and one student, Louise insists and demands an answer to the question by gesticulating with his hand, and saying: “But...” (Turn 19). However, the teacher does not answer Louise; instead she proceeds with the lesson and finishes the dialogue in this way: “We leave the transition elements and focus on group III A (...) all the atoms strive towards getting 8 electrons in the outer shell...” (Turn 20).

Transcript D: Students were not satisfied with teacher explanations; they are disappointed with the teacher because of the limitation placed upon their own inquiry. The linear and sequential character of science learning and assessment is exposed.

Main dialogue				Parallel dialogue	
21	18.25	Emil	Palladium has 18 electrons in the outer shell!!! [Paladium har 18 elektroner i det ytterska skal!]		
22	18.26	Teacher: Maria	Do not get yourself entangled with transition elements!!! (...) they behave silly, I do not want to talk about them!! They are anomalous! Forget them! ... Until you begin to study chemistry at university!! [The teacher becomes irritated] [Trassla inte in med övergångs element!!! De bete sig dumma, jag vill inte prata om dem!! De är oregelbundna! Glöm de!... Tills ni börjar studera kemi på universitetet!!!]	Charles	Hanna! do you understand? [Hanna! Förstår du?]
23	18.27			Hanna	Whats?[Va?]
24	18.28			Charles	Do you understand? [Förstår du?]
25	18.29			Hanna	I think I have missed something. [Jag tror jag har missat något]

				Main dialogue	Parallel dialogue
26	18.31				Charles What? [Vad?]
27	18.32				Hanna I don't understand.[Jag förstår inte]
28	18.33	Jannet	Oh my God!	Charles	It is those transitions...[de är de övergångs element].
29	18.34			Jon	What's? [Vad?]
30	18.35			Charles	It is really unclear. [Det är verkligen oklar]. (Oct 2009)

Emil breaks out of the IRE pattern, stating that “Pd has 18 electrons in the outer shell!!!” (Turn 21). The teacher becomes angry and she silences the student by appealing to the authority of her position (Carlsen 1997) saying: “...Do not get yourself entangled with transition elements (...)” (Turn 22). She again is appealing to her authority of knowledge and says “they are anomalous” (Turn 22). Eventually, the teacher moves from a scenario of openness to a scenario of power over meaning and said: “Forget them! ...” (Turn 22). The teacher finishes the dialogue by revealing her conception of learning, which she presents as linear and sequential: “Until you begin to study chemistry at university!! [The teacher becomes irritated]” (Turn 22). This interaction suggests that the teacher’s assumption of learning as linear and sequential also informs the assessment model she endorses, which may be relevant when she grades her students. Such an approach also may foster teaching and assessment practices that focus on specific components of separate and hierarchical skills; a strategy that may be less suitable if the aim of instruction, as the national Swedish curriculum stipulates, is to develop a learner’s thinking or problem-solving abilities.

The dialogues between students Charles, Hanna, and Jon provide confirmation of their engagement on learning the subject, and shared goals for learning but they did not understand the transition elements and they did not dare to ask the teacher; instead they sought to involve members of the class in a side discussion. Charles said: “Hanna, do you understand?” (Turn 22). Charles asked again, “Do you understand?” (Turn 24). Charles turned to Jon in an attempt to gain an understanding of the subject: “It is those transitions...” (Turn 28). Eventually, Charles responded to Jon: “It is really unclear...” (Turn 30). The students had learned from prior experience that to interrupt the teacher’s lecture is not acceptable and as such the teacher is the only one with the authority to talk (Lemke 1990).

What the teachers said about their assessment practices in everyday classroom activities

The following is a conversation with the teachers Maria and Susanne who were asked to answer particular questions about their assessment practices.

Interviewer

What is your conception of the students’ knowledge of the subject when they ask questions during your lecture, and how do you assess students when they ask questions?

Response by Teacher Maria

It is individual. You notice that there are students who ask because they think and are keen to ask questions, but there are other students who are more annoying since one notices

that they can't follow the lecture; they don't listen and they ask stupid questions (...). I try not to assess students who ask stupid questions, but for those who ask good questions I feel that this should be included in the total grading of the course. But I try not to lower the grades of those who ask stupid questions

Interviewer

During the lessons, why are all the students silent? Why do the students not talk at all? They do not take the initiative and participate in discussions—why?

Response by Teacher Maria

Ha, ha, I do not give them a chance to participate! (...) It's clear that I should ask them a lot more, perhaps! (...) But actually I want the class to be quiet (Nov 2009).

Her comments reveal that Maria wants students to be quiet and silent: “Ha, ha, I do not give them a chance to participate! (...) actually I want the class to be quiet.” The transcript also shows that she has a preconception about students' characteristics: “but there are other students who are more annoying (...), and ask stupid questions.” This pattern of interaction is relevant because it reveals the social and cultural character of learning and assessment. Such practices seem to be deep-rooted in the behavior of the teachers suggesting that they consider their practices to be normalized and approved by the school. Both her personal experiences and professional education are culturally conditioned and these values and conceptions on learning may affect inevitably her conceptions on assessment. The assessment of students is related to the classroom context in which she constructs a culture of assessment based on socially accepted norms. The transcript of the conversation with the chemistry teacher Maria confirms that the teacher controls the talk when teaching chemistry.

Conversation with the biology teacher, Susanne:

Interviewer

What is your conception of the students' knowledge of the subject when they ask you questions during your lecture and how do you assess the students when they ask questions?

Teacher Susanne

If I have to repeat things I think the students should know already that I become a little... you can't say that I give them a negative assessment... but I feel suddenly that this student is perhaps not one of the top students. If it's a student who asks questions and is active and is asking questions, then I can assess the student more positively... That's how it goes! It is also important when you assess students that they feel comfortable; it has to do with personality—there are many things to consider when grading. But it is the case that you grade up or down depending on what kind of questions they ask

Interviewer

During the lessons, why are all the students in silence? Why do the students not talk at all? They do not take the initiative and participate in discussions—why?

Teacher Susanne

It's force of habit, as I come from a university teaching style. But it is the students who want it this way. The students want me to tell them all the important facts in the subject and therefore I must talk all the time (Nov 2009).

The transcript shows that the teacher is more comfortable with lectures governed by recitation, in an IRE pattern, and she admits to having silenced students: “It's force of

habit” but one that she justifies by claiming that the students also want the class to be this way. The transcript also shows that the teacher assesses the students’ questions: “But it is the case that you grade up or down depending on what kind of questions they ask.” She categorizes students as “top” students and she maintains that she assesses students’ questioning as being positive or negative: “if it’s a student who asks questions and is active and is asking questions then I can assess the student more positively.” She claims that a top student is evaluated more positively, saying: “...That’s how it goes!” In addition, feelings, personality, and other unspecified things are important when this teacher grades students: “(...) they feel comfortable; it has to do with personality—there are many things to consider when grading.” (Nov 2009). This claim suggests that the teacher takes account of the characteristics of a student’s personality when she assigns a grade.

A summary of the teachers’ statements show that both teachers have similar conceptions about learning and assessment. Both Maria and Susanne’s transcripts show how these teachers have assessment practices that are at odds with established measurement principles, that is, when assessing and grading students they do not take into account personal characteristics of students.

In the pedagogical discourse of IRE it is still the teacher who decides who will talk, and when, and about what (Lemke 1990). Susanne blamed students for her own instruction practices: (...). “But it is the students who want it this way.” She might be advised to explicitly counteract the reproduction of instructional discourses that coerce students. We suggest that an effort to do this should entail having a more explicit awareness of the social positioning of teachers as individuals with possibilities for creating and sharing their own reflections with their colleagues as to how their actions shut down students.

Let us explain the sense we made

A summary of the discourses displayed in the in-depth conversations with the teachers revealed that all teachers used implicit assessment of their students in the everyday school context. The teachers, Susanne and Maria, recognized that they used these informal assessments to grade students in circumstances where students ask good questions and are active in discussions. Maria claimed that when she assessed students who posed good questions, this kind of assessment should influence the final grading of students. However, when students posed “stupid questions” Maria tried not to assess them. This statement may be understood in the light of what is seen as socially reasonable as well as what one is expected to say in different situations. According to convention or “common-sense” assumptions, it is not acceptable to say openly that students should be evaluated negatively in such situations. Teacher Susanne shared this practice with her colleague Maria and evaluated “top students” more positively; however, differently from Maria, she evaluated “up and down” depending on the questions students posed. Susanne also shared the lecturing approach with the teacher Maria and, similar to Maria, she also recognized that she silences students, claiming that it is because of “habit.” Both teachers shared and practiced “common-sense” assumptions in their everyday work according to which they interacted and constructed linguistically how the assessment proceeded during the social interactions. Clearly, assessment may well be considered not only as a technical endeavor but also as culturally conditioned. Thus, classroom practice may to some extent influence assessment and grading. Historically, however, mainstream educational assessment researchers examining teachers’ cognition may have ignored contextual variables that lead us to ignore teachers’ thinking in specific contexts.

We have found that the assessments teachers perform have both a formal character, particularly when using high-stakes testing and examinations, and an informal character that involves assessment of student's personal characteristics in everyday interactions in classroom, which in turn the teachers use for grading purposes. We consider that both the formal and informal approaches of assessment are different from our meaning of the same, given the way they are enacted in practice. Both approaches are nevertheless similar as both have same purpose. Let us explain.

Formal assessment specialists are very clear that individuals' characteristics should not be taken into account when students' achievement is assessed. However, in this kind of assessment the teachers follow established conventions, which were stipulated in the curriculum, both in turn depending on social rules, institutions and conventions in society. In this context the formal assessment is supposed to be properly performed, justified and objective. Informal assessment, in contrast, often involves teachers taking into consideration students' personal situations and characteristics. However, both the formal and informal assessments were done at the same time, and both led the teacher to make decisions about grades, that in turn were used to make decisions about students' lives. Doing this does not allow the teacher to differentiate between formal and informal but encourages the teacher to allow the formal to override the informal.

We argue that using formal and informal assessments simultaneously is a strategy teachers use to screen students in a tradition that only the "right" kind of students may be prepared for university science studies. Teachers' instruction is characterized by a dominance of the science discourse leaving students alone with their own reflections. The discourses of the teachers are imbued in a pedagogic discourse (Bernstein 1996) with the purposes to legitimize the schools' authoritative regulations of order, relations and identity. The teachers follow how the Swedish Ministry of Education suggests that they teach science. Accordingly, in order to achieve their mission, teachers take charge of classroom by exercising strong control over both the learning context and what they view as science knowledge. However, the context created by the teachers is to privilege the needs and abilities of some students, while simultaneously ignoring the needs of others. Classroom assessment can achieve the fundamental goals of differentiation, evaluating not only disciplinary content knowledge but also critical skills necessary for social advancement. Furthermore, we hold that teacher-generated classroom assessments meet the needs of both formal and informal assessments. Research on informal assessment is advancing, and we reflect on the importance of linking informal assessments with learning process involving authentic inquiry is (e. g. Meyer and Crawford 2011).

Our final thoughts

We have shown in this study a problematic aspect of discourses in instructional contexts that constrain the learning and development of students. We are of the opinion that one way to overcome this issue is to change the classroom practice though it is easier to say than to do it. Such practices seem to be ingrained in the behavior of the teachers suggesting that they consider their practices to be normal, endorsed by schools. The observed classroom interactions including teachers shutting down specific students. This resulted in excluding students from exercising their learning needs with regard to science thereby hindering their development of scientific literacy. We do not consider the solutions are straightforward because changes to norms and habits carry contradictions between people as the established conventions created by order of discourses may be undermined by new

orders of discourse. We want here to draw attention to the idea, promoted by Fairclough (2003), that discourses involve individuals and require individual creativity, but the outcome of restructuring orders of discourse lies outside individuals in social competition with each other. Participants' agency in restructuring orders of discourse involves the individuals' will—or lack thereof—to mediate in the world. For example, an individual who looks to persuade others through their discursive analysis demonstrates strong levels of agency whereas others who are passive demonstrate weak levels of agency (Roth 2007).

Further research may encourage teachers to interact with their students in a fashion, which develops explicitly and purposefully their students' command of both conversational and academic language. These changed interactions consequently will improve the quality of their students' learning, their performance on science assessments, and—we hold importantly—their development as social agents. To this end, educational research can and should emphasize not only the epistemic and technical aspects of assessments, but also the social and cultural aspects thereof.

We were limited in certain areas

Research on social interaction shows that what is observed at a particular moment is not just a function of the current interactions, but the deployment of resources that have been shaped and renewed in many other prior interactions (Fairclough 2003). Consequently, any analysis of discourse is selective because we choose to examine specific questions about certain social events and not other conceivable matters. But as Fairclough (2003) shows deeper cultural norms that have become so accepted by the participants, like the teachers interviewed in this study and the students struggling to learn, are not readily open to critical analysis. A particular limitation of our study is the lack of focus on the temporal dimensions of learning in the studied context (Mercer 2004). The following quotation from biology teacher Susanne, extracted from the video-recorded lectures, shows a typical example of such temporal dimensions in classroom interactions.

Susanne: Tomorrow we will visit the laboratory MCT. As I told you last week, this laboratory is working with ... It is located in (...). You can read more about this activity on your homepage (...). We will meet at the parking place outside the building (...). [Imorgon ska vi besöka labbet MCT. Som jag berättade förra veckan, på denna labb arbetar man med ... Den ligger i ... Ni kan läsa mer om platsen och aktiviteten på Er sida (...). Vi träffas vid p-platsen utanför byggnaden] (Sep 09).

This kind of discourse shows the past shared experiences that took place before the day of our ethnographic observation, as well as activities to be realized outside the classroom and in the future. These activities involve teaching, learning, and probably assessment processes and are examples of activities that have not been the focus of our study. Besides, it is not possible to capture accurately events involving informal assessment by teachers—these are only understood through teachers' statements in the interviews. All dimensions of the social life created, the time and place, and other categories in the day-to-day interactions among members of the studied groups are socioculturally constructed and it is not possible to describe them in totality, nor is it possible to use the findings alone as a basis for inferring something about what the members of the studied group are doing (Fairclough 2003).

References

- Beach, D. (1999). Om demokrati, reproduktion och förnyelse I dagens gymnasieskola. [About democracy, reproduction and renewal in the days Upper-secondary school]. *Pedagogisk Forskning i Sverige*, [Pedagogic research in Sweden], 4, 349–365.
- Bernstein, B. (1996). *Pedagogy, symbolic control and identity*. London: Taylor and Francis.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–71. doi:[10.1080/0969595980050102](https://doi.org/10.1080/0969595980050102).
- Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. doi:[10.1007/s11092-008-9068-5](https://doi.org/10.1007/s11092-008-9068-5).
- Broadfoot, P. M. (1996). *Education, assessment and society*. Philadelphia: Open University Press Buckingham.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education*, 11(1), 7–27. doi:[10.1080/0969594042000208976](https://doi.org/10.1080/0969594042000208976).
- Broady, D., Andersson, M. B., Börjesson, M., Gustafsson, J., Hultqvist, E., & Palme, M. (2000). Skolan under 1990-talet- Sociala förutsättningar och utbildningsstrategier. In: SOU2000:39 (Ed.), *Statens Offentliga Utredningar* (pp. 2–133). Stockholm, Sweden: Statens Offentliga Utredningar.
- Carlsen, W. S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research*, 61, 157–178. doi:[10.3102/00346543061002157](https://doi.org/10.3102/00346543061002157).
- Carlsen, W. S. (1997). Never ask a question if you don't know the answer: The tension in teaching between modeling scientific argument and maintaining law and order. *Journal of Classroom Interaction*, 32(2), 14–23.
- Dahlstedt, M. (2007). I val(o)frihetens spår: Segregation, differentiering och två decennier av skolreformer. [The consequences of (un)free choice policies: Segregation, differential assessment and educational reform in Sweden since the 1980th]. *Pedagogisk Forskning I Sverige*, 12(1), 20–38. [Pedagogic research in Sweden]. Stockholm.
- Delandshere, G. (2002). Assessment as inquiry. *Teachers College Record*, 104, 1461–1484. doi:[10.1111/1467-9620.00210](https://doi.org/10.1111/1467-9620.00210).
- Fairclough, N. (2001). *Language and power* (2nd ed.). London: Pearson Education Limited.
- Fairclough, N. (2003). *Analysing discourse. Textual analysis for social research*. London, NY: Routledge.
- Forsberg, E. & Lindberg, V. (2010). *Svensk forskning om bedömning—En kartläggning*. Vetenskapsrådet. Vetenskapsrådet rapportserie, 2:2010. Stockholm. [The Swedish research on assessment—a Mapping. The Swedish Council of Science Research 2:2010, 1–135. Stockholm].
- Gallard M., A. J., Pitts, W., Flores Bustos, B., Claeys, L., Baynes, G., Brkich, K., Stevenson, A., & Wassell, B. (2013, November) *Latina resiliency: In pursuit of STEM careers*. Paper presented at the meeting of the World Education Research Association (WERA), Guanajuato, Mexico.
- Gallard Martínez, A. J., & Antrop-González, R. (2013). The use of de jure to maintain a de facto status quo. *Cultural Studies of Science Education*, 3, 991–996. doi:[10.1007/s11422-013-9523-6](https://doi.org/10.1007/s11422-013-9523-6).
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355–392.
- Gómez, M., & Jakobsson, A. (2013). *Classroom assessment, grading practices and teachers' epistemological beliefs on science learning in upper secondary schools* (in press).
- Goodall, H. L, Jr. (2000). *Writing the new ethnography*. Maryland: AltaMira Press, Rowman & Littlefield Publishers.
- Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice* (3rd ed.). London, NY: Routledge.
- Harlen, W. (2008). Trusting teachers' judgement. In S. Swaffield (Ed.), *Unlocking assessment: Understanding for reflection and application* (pp. 138–153). New York: Routledge.
- Klapp Lekholm, A. (2010). Vad mäter betygen? [What measures grades?] In C. Lundahl & M. Folke-Fichtelius (Eds.), *Bedömning i och av skolan—praktik, principer, politik* [Assessment in and of school: Practice, principle, policy] (pp. 129–142). Lund: Studentlitteratur.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press. doi:[10.1017/CBO9780511815355](https://doi.org/10.1017/CBO9780511815355).
- Lemke, J. (1990). *Talking science: Language, learning, and values*. Westport, Connecticut: Ablex Publishing Corporation.
- Mack, L. (2012). Does every student have a voice? Critical action research on equitable classroom participation practices. *Language Teaching Research*, 16, 417–434. doi:[10.1177/1362168812436922](https://doi.org/10.1177/1362168812436922).
- Martin, A. M., & Hand, B. (2009). Factors affecting the implementation of argumentation in the elementary science classroom: A longitudinal case study. *Research in Science Education*, 39, 17–38. doi:[10.1007/s11165-007-9072-7](https://doi.org/10.1007/s11165-007-9072-7).

- Mehan, H. (1979). *Learning lessons, social organization in the classroom*. Cambridge: Harvard University Press. doi:[10.4159/harvard.9780674420106](https://doi.org/10.4159/harvard.9780674420106).
- Mercer, N. (2004). Sociocultural discourse analysis: Analyzing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, 1, 137–168. doi:[10.1558/japl.2004.1.2.137](https://doi.org/10.1558/japl.2004.1.2.137).
- Meyer, X., & Crawford, B. A. (2011). Teaching science as a cultural way of knowing: Merging authentic inquiry, nature of science, and multicultural strategies. *Cultural Studies of Science Education*, 6, 525–547. doi:[10.1007/s11422-011-9318-6](https://doi.org/10.1007/s11422-011-9318-6).
- Mortimer, F. E., & Scott, P. (2003). *Meaning making in secondary science classrooms*. Philadelphia: Open University Press.
- Nystrand, M., Gamoran, A., Kachur, R., & Prendergast, C. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Columbia University, Teachers College Press.
- Nyström, E. (2007). *Talking and taking positions. An encounter between action research and the gendered and racialised discourses of school science*. Unpublished doctoral dissertation, Umeå University, Umeå, Sweden.
- Robson, C. (2007). *Real world research: A resource for social scientists and practitioner-researchers* (2nd ed.). Oxford: Blackwell Publishing.
- Roth, W.-M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content, and student responses. *Journal of Research in Science Teaching*, 33, 709–736. doi:[10.1002/\(SICI\)1098-2736\(199609\)33:7<709::AID-TEA2>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1098-2736(199609)33:7<709::AID-TEA2>3.0.CO;2-R).
- Roth, W.-M. (2007). Theorizing passivity. *Cultural Studies in Science Education*, 2, 1–8. doi:[10.1007/s11422-006-9045-6](https://doi.org/10.1007/s11422-006-9045-6).
- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Los Angeles: Sage.
- Skolverket. (1998). *Regeringsuppdrag "Konsekvenser av de nya behörighetsreglerna till gymnasieskolan" – slutrapport* (regeringsbeslut 49, 1998-12-17). [Report of The Swedish National Agency for Education 1998:49].
- Skolverket. (2013). *The Swedish curriculum for the Upper-secondary school* [The Swedish National Agency for Education 2013]. Retrieved from www.skolverket.se.
- SOU. (2000). En uthållig demokrati. Slutbetänkande. Statens Offentliga Utredningar 2000:1 [A sustainable democracy. Final report. The Swedish Government Official Report, 2000:1].
- SOU. (2004). Tre vägar till den öppna högskolan. Statens Offentliga Utredningar, 2004:29 [Schools for Open High Education. The Swedish Government Official Report, 2004:29].
- Thornberg, R. (2010). School democratic meetings: Pupil control discourse in disguise. *Teaching and Teacher Education*, 26, 924–932. doi:[10.1016/j.tate.2009.10.033](https://doi.org/10.1016/j.tate.2009.10.033).
- Tobin, K., & Gallagher, J. J. (1987). The role of students in the science classroom. *Journal of Research in Science Teaching*, 24, 1–75. doi:[10.1002/tea.3660240107](https://doi.org/10.1002/tea.3660240107).
- Utbildningsdepartementet. (1994). *Läroplaner för det obligatoriska skolväsendet och det frivilliga skolformerna*. Lpo 94-Lpf 94. Ministry of Education, Curriculum for compulsory school system and voluntary school forms. Stockholm.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds. and Trans.), Cambridge, MA: Harvard University Press.
- Wells, G. (1993). Reevaluating the IRF sequence: A proposal for the articulation of theories of activity and discourse for the analysis of teaching and learning in the classroom. *Linguistics and Education*, 5, 1–37. doi:[10.1016/S0898-5898\(05\)80001-4](https://doi.org/10.1016/S0898-5898(05)80001-4).
- Wertsch, J. V. (1991). *Voices of the main: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.
- Wikström, C. (2005). Grade stability in a criterion-referenced grading system: The Swedish example. *Assessment in Education*, 12, 125–144. doi:[10.1080/09695940500143811](https://doi.org/10.1080/09695940500143811).
- Wikström, C., & Wikström, M. (2005). Grade inflation and school competition: An empirical analysis based on the Swedish upper secondary schools. *Economics of Education Review*, 24, 309–322. doi:[10.1016/j.econedurev.2004.04.010](https://doi.org/10.1016/j.econedurev.2004.04.010).

María del Carmen Gómez is a PhD student in the department of Teacher Education in Malmö University in Sweden. Her research is based on exploration of teachers' assessment practices in science education in Sweden and her work is grounded in sociocultural theories of learning and assessment. She believes strongly

that natural sciences should be available for all students and that assessment should be used to enhance learning.

Anders Jakobsson is professor in science education at the faculty of Learning and society at Malmö University in Sweden. His research is focused on language and discursive perspectives on learning in science and on issues related to equality in science education. In his research he seeks to understand the tendencies and trends in students' science knowledge and epistemology and to explore the connection between the results of large-scale studies and how science teaching and learning are constituted in diverse school environments.

Paper III



Student explanations of their science teachers' assessments, grading practices and how they learn science

María del Carmen Gomez¹

Received: 18 January 2016 / Accepted: 11 March 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The current paper draws on data generated through group interviews with students who were involved in a larger ethnographic research project performed in three science classrooms. The purpose of the study from which this data was generated, was to understand science teachers' assessment practices in an upper-secondary school in Sweden. During group interviews students were asked about their conceptions of what were the assessment priority of teachers, why the students were silent during lecturing and their experiences regarding peer- and self-assessments. The research design and analysis of the findings derives from what students told us about their assessments and learning sciences experiences. Students related that besides the results of the written test, they do not know what else teachers assessed and used to determine their grades. It was also found that students did not participate in the discussion on science because of peer-pressure and a fear of disappointing their peers. Student silence is also linked with student conceptions of science learning and student experiences with methodologies of teaching and learning sciences.

Keywords Student perspectives · Teacher's assessment · Science learning · Swedish upper-secondary school

Sammanfattning Den här artikeln bygger på data som kommer från gruppintervjuer med elever som deltagit i ett större etnografiskt forskningsprojekt utfört i tre klassrum på de naturvetenskapliga programmen. Syftet med studien var att förstå lärares bedömningspraktiker i de naturvetenskapliga programmen i en gymnasieskola i Sverige. Under gruppintervjuer frågades eleverna om deras uppfattningar om vad lärare prioriterar när de bedömer elevernas lärande, varför eleverna var tysta under lektionerna och deras

Lead Editor: Alejandro J. Gallard-M.

✉ María del Carmen Gomez
maricarmen7879@gmail.com

¹ Faculty of Social Sciences, Lund University, Lund, Sweden

erfarenheter när det gäller kamrat- och självbedömning. Forskningsdesign och analys av resultaten härrör från vad eleverna berättade om sitt lärande och deras bedömningserfarenheter. Studenter berättade att, förutom resultatet av det skriftliga provet, vet de inte vad lärare bedömer och vad lärare använder för att bestämma deras betyg. Det visade sig också att eleverna inte deltar i diskussioner om vetenskap, på grund av gruppträck och en rädsla för att göra sina kamrater besvikna. Studenternas tystnad hänger också samman med studenternas föreställningar om naturvetenskapligt lärande och studenternas erfarenheter av metoder för undervisning och lärande av naturvetenskapen.

Nyckelord Elevperspektiv · Lärarens bedömning · Vetenskap lärande · Svenska gymnasieskolan

The grounding of my thinking

From a pedagogical standpoint, learning requires student participation, and assessment should support learning provided that the assessment promotes learning. Research in assessment suggests that assessment has to move from the *summative* character to the *formative* character (e.g., Torrance 2007), whereby assessment procedures and practices are developed to support learning (e.g. Black and Wiliam 2009). In formative assessment, key elements of assessment as self- and peer- assessment require students to participate in the process and, in collaboration with teachers, come to a more meaningful learning experience. The idea of more meaningful learning experiences raises questions regarding the conditions necessary, which would enable changes to take place within schools.

In Sweden there is a decades-long tradition, of promoting student participation in their education:

The school has the obligation to give students increased responsibility and empowerment in line with their age and maturity. (Stipulation 1980:64) [*Skolan har skyldighet att ge eleverna ökat ansvar och medinflytande i takt med deras stigande ålder och mognad.*] [Förordning 1980:64]

The 1994 curriculum, Lpf 94, confirmed students' responsibility and participation in fundamental educational legislation "Right and Obligations" that has been endorsed in the last revised national curriculum 2011:

It is not in itself sufficient that education imparts knowledge of fundamental democratic values. It must also be carried out using democratic working methods and develop the students' ability and willingness to take personal responsibility and participate actively in societal life. (Swedish Ministry of Education 2013, p. 6)

The Swedish curriculum is process oriented, based on modern Western theories of learning. For example, as regarding the concept of knowledge it states that:

Knowledge is a complex, multi-faceted concept. Knowledge can be expressed in a variety of forms – as facts, understanding, skills, and familiarity and accumulated experience – all of which presuppose and interact with each other. Teaching should not emphasize one aspect of knowledge at the cost of another.

Students' acquisition of knowledge is dependent on developing the ability to see interconnections. The school should enable students to gain a general and coherent

view. Students should get the opportunity to reflect over their experiences and apply their knowledge. (Swedish Ministry of Education 2013, p. 6)

Consequently, Swedish teachers are requested to assess student's total abilities:

Teachers when awarding grades should use all the information available about the student's knowledge in relation to the national knowledge requirements for each course, also take into account the knowledge that a student has acquired outside actual teaching, and on the basis of the national knowledge requirements for each course make an all-round assessment of each student's knowledge. (Swedish Ministry of Education 2013, p. 13)

In fundamental values and task for the school the curriculum supports cooperative perspective in teaching and learning:

Students should develop their ability to take initiatives and responsibility, and to work both independently and together with others. (Swedish Ministry of Education 2013, p. 6)

It is the school, which must guarantee students the opportunity to exercise their rights, and influence their education:

Opportunities to exercise influence over their education and take responsibility for their studies requires that the school clarifies the goals of education, its contents and working forms, as well as the rights and obligations that students have. (Swedish Ministry of Education 2013, p. 5)

The use of formative characteristics of assessment as for example self-assessment, is also acknowledged in the revised national curriculum of 2011, which stipulates that:

The goals of the school are that all students individually: take responsibility for their learning and study results and can assess their study results and need for development in relation to the requirements of the education (Swedish Ministry of Education 2013, p. 13)

Regarding student work, the newly revised curriculum of 2011 stipulates that:

Students should develop their ability to think critically, examine facts and relationships, and appreciate the consequences of different alternatives. By these means students will come closer to scientific ways of thinking and working (Swedish Ministry of Education 2013, p. 5)

Clearly, the Swedish national curriculum has been intended to advance a learning situation where students take an active and responsible part, and are assessed in a variety of areas with different methods, supporting the learning. Accordingly, I am trying to understand students' experiences in their participation in learning and assessment processes as well as their conceptions of teachers' assessments. Observed contradictions to current and previous national curricula indicate that student participation in the discourse of science is somewhat limited and that students do not participate in their assessment of learning (e.g., Engström and Carlhed 2014).

This position is consistent with the finding of an ethnographic study in which it was found that the students sat silent in the classroom during the teacher's lesson (Gómez and Jakobsson 2014). Specifically, students answered teachers' questions posed in a pattern known as Initiation–Response–Evaluation (IRE). Even though teachers tried to engage

students by asking questions students not actively involved in their learning by using this method. Specifically, the new curricula require that students participate in the learning process and that teachers promote student participation in classroom. However, when teachers try to engage students using IRE the questions asked are from the teachers and further questioning is not encouraged from the students.

The noted discrepancy between aim of the national Swedish curriculum and student response, challenged us to find out why students are silent during the lessons. Also, the ethnographic study revealed that everyday activity in the classroom was seemingly characterized by written examinations and assignments, in contradiction with the curricula. I wondered what other kinds of assessment the students had experience with, besides written examinations.

My main research idea and the research questions

Thus, the purpose of the current study is to examine student experiences and their participation in learning and assessment processes as well as their conceptions of teacher assessments. My research questions are as follows:

1. What are the students' conceptions of teachers' assessments?
2. What is particularly important for the teachers to assess above all for the purpose of grading as understood by students?
3. Do students know what teachers assess, in addition to the test and assignments?
4. What are student's experiences of peer- and self- assessment?
5. Why are students silent most of the time while the teachers are lecturing?

Assessment is important from a democratic point of view, for the individual and not least for the society. We are living in a meritocratic society, where individuals should be allowed to compete on an equal basis depending on individuals' different merits. The ways to investigate and assess individuals has major implications for students, as have the methods, which are used for assessment. In Sweden it is the teachers who have the sole prerogative to assess students' outcomes. Therefore, it is important to understand how teachers are assessing the students. My main idea is to analyse and to further understand the collaboration between students and teachers regarding educational assessment and learning sciences, by analysing students' conceptions and experiences with teacher assessments. I consider it important to enable the engagement of students in matters, which concern them by inducing some reflections about how they learn sciences and how teachers assess them. The framework I used is the Swedish national school system is based on the democratic foundation of fostering a democratic education for all citizens. In addition, the framework also requires open democratic debate, which reflects the scientific method, comprising, amongst other things, an open, respectful discussion containing clear and transparent arguments.

In what premises is my research grounded?

During the observation period, except for the teachers' voices, the silence was absolute in the classroom during the chemistry lessons I observed. I realized that the observations on teacher assessment in the classroom only give us clues about how students perceive

learning and assessment and not their participation during a science lesson. I observed in this ethnographic study different kinds of examinations to include paper-and-pencil in all observed science classes and at all the studied levels of upper secondary school. I understood that students took notes during the lectures and often whispered to each other with the aim of understanding issues and solving problems by themselves, unobtrusively. Sometimes someone interrupted the teacher and asked about something they did not understand, although the general pattern observed was that students would sit in silence most of the time, while the teacher conducted the lecture. I became curious during the observation period as to how students perceive their own silence in relation to learning the science content. Is there some kind of hidden "rule", that everyone should be quiet and which, they have internalized, in some way? Accordingly, I considered it important to listen to students and ask them the reason(s) why they were silent during the lecture.

I considered it important to ask students regarding any additional assessment methods in use, as may be expected from the curricula instruction because we wanted to know, how do students participate in assessment of their learning? Specifically, I was curious as to how students are involved in assessment and, if they are, what experiences did they have that helped them understand the characteristics of the assessment, such as whether it was formative or summative, or whether it was constructive and covered knowledge attained outside school?

What is the research on assessment in an educational context telling us?

Several shifts in educational assessment theories and practices have described many different changes and have given many different terms for assessment defined more or less differently (e.g., Newton 2007). Those changes have caused some confusion about the term *assessment*, both worldwide and in Sweden. Much of the confusion may have been caused by a lack of common agreement on how to define an exact meaning of the concepts involved when used in an educational context, particularly regarding information on student achievement and general outcomes in education. An example of confusion is the use of the terms *assessment* and *evaluation*, which according to Anders Lysne (2006), have been used synonymously in spite of it being possible to draw a distinction between them.

Evaluation is connected to judgement regarding goodness, worth and values being influenced by cultural values and general purposes of education. The concept of assessment is merely used in connection with testing and judging potentials of intellectual recourses, attainments and performances, mostly at the individual level. Gregory Cizek (2000) suggests that both the concepts of assessment and of evaluation refer to the planned observation and gathering of information, even if *assessment* may be slightly broader in scope, involving several tests or other sources of information. He argues that the range of the term *evaluation* is narrower, and usually comprehends assessment, and that *evaluation* implies the process of ascribing merit, or value, to the results of an observation or data collection. Paul Newton (2007) describes three levels of assessment purposes and explains that an assessment system, which is fit for one purpose, will not necessarily be fit for all purposes. These levels are: (a) the *judgement* level, involving the technical aim of an assessment event, as for example the purpose to obtain standard-referenced judgement, expressed as a grade, and representing an assessed person's knowledge, skill or understanding, i.e. the account of their competence; (b) the *decision* level, concerns the use of an assessment or judgment, the decision action or process which enables it to support a

selection for entry to higher education. According to Newton (2007), this is the most significant usage of the term “assessment purpose” (p. 150) because this level, seems to be the level that is most frequently associated with research assessment; and (c) the *impact* level. Newton explains that, in order to ensure that students remain motivated; assessment at the impact level requires that all the students learn a subject’s common core. For example, the assessment might be administered on a unit-by-unit basis and aligned to a national curriculum.

In this article, when I refer to assessment it is those actions, which teachers perform at the *decision level*, and when I refer to grades I am referring to the *judgement level*; when I refer to assessment for learning or assessment for formative purposes I am referring to assessment at the *impact level*.

Assessment for formative purposes has different characteristics from that used summatively, that are for accountability purposes (Wiliam and Black 1996). An assessment, which serves formative functions, “elicits evidence that yields construct-referenced interpretations that form the basis for successful action in improving performance, whereas summative functions prioritize the consistency of meaning across contexts and individuals” (Wiliam and Black 1996, p. 537). Important aspects of formative assessment are the relationship between teacher and learners and the activation of students “as the owners of their own learning” and “Activating students as instructional resources for one another” (Black and Wiliam 2009, p. 8). Often it is the teacher who collects the evidence of learning and makes the decisions preceding grading. In formative assessment it is important that learners participate in the grading decision.

Since the responsibility for learning stands not only with teachers but also learners, it is not surprising that an important element of formative assessment is peer- and self-assessment (Black and Wiliam 2009). When students are working with their peers the aim is to raise achievement (Harrison 2011). Peer-to-peer assistance can be called collaborative work, cooperative learning, peer tutoring, peer-based assessment or peer-directed learning. However, the foundation of these methods has a historical origin based on learning. “Cooperative learning exists when students work together to accomplish shared learning goals” (Slavin et al. 2003, p. 177).

Research on different collaborative learning approaches shows that, although there is consensus on the advantages of collaborative learning, the issue is the subject of controversial discussions because of the different approaches adopted. For example, Christine Harrison (2011) claims that placing students in groups to work collaboratively does not always mean that learning has taken place. This is, according to her, because what is important in cooperative learning is the nature of interactions between students and the social characteristics of the classes. Peer-to-peer assessment is a complex activity, which demands knowledge, organisation and training.

In practice, peer-assessment is an important stimulus to self-assessment. When students help their peers, they have to master the subject themselves to teach others; they learn by taking the roles of teachers and assessors of others (Sadler 1998). Self-assessment is also inherent to learning because the individuals have to be involved in judging whether or not they have reached the learning goals. According to Harrison (2011) peer- and self-assessment play a major role in how students learn and how teachers teach and that it is not evident in all classrooms. Where it does appear, teachers have usually made a pedagogical decision as to its importance in learning, since it requires time, planning and skill to implement successfully. The teachers’ role in this activity is critical in developing, planning and supporting students’ engagement with one another in the creation of collaborative work and dialogues. In terms of grouping interaction, Peter Blatchford, Ed Baines,

Christine Rubie-Davies, Paul Bassett and Anne Chowne (2006) argue that peer-collaboration can be successfully adopted in the everyday classroom work, if teachers have trained the students.

What does science education look like in the Nordic countries?

Educational research in Nordic countries shows that both teaching and assessment are activities strongly controlled by teachers and that these activities are teacher-centred (e.g., Lavonen, Angell, Bymen, Henriksen and Koponen 2007). For example, an evaluation of schools, performed in upper-secondary schools in Denmark (Norriild, Angell, Bang, Larsen, Paulsen and Stubgaard 2001), found patterns of teaching in physics that were teacher-centred. In this research, teacher lecturing is described as being related to students' active participation in classroom talk, and that these teacher–student interactions have an assessment character. Similar patterns are found in an evaluation of the current Norwegian national curriculum for grades 1–10. It was found that the lessons are teacher-centered, and that students experience sciences as a theoretical subject and that written examinations are the most common assessment method used by teachers (Haug 2003).

Research on social interactions in upper-secondary physics classes in Finland (Lavonen, Angell, Bymen, Henriksen and Koponen 2007) found that teacher-delivered instruction is criticized by students because they are not involved in this activity, and that this kind of instruction often fails to challenge students to reason at higher levels of thinking, analysis and evaluation. Students preferred more discussions and argumentation activities in the classroom. Analysis of discourses of school science in upper-secondary schools in Sweden, by Eva Nyström (2007), shows that science teachers and science students have a deeply rooted tradition. The tradition involves notions of status that give science teachers a strong voice in the school and creates expectations of high performance among science students. She suggests that science is institutionally constructed as difficult and complex and that science teaching needs teachers with the highest possible intellectual capacity. Inequality and power tensions have also been revealed in the science classroom at upper secondary school by Nyström (2007) who analyzed teacher talk she describes as long established hierarchies and taken-for-granted values of science subjects that promote the reproduction of male gender. As a result students, depending on their gender, are positioned outside or inside the discourses. She describes how students and teachers ascribe meaning to gender, social class and ethnicity in different circumstances.

Susanne Engström and Carina Carlhed (2014) have studied different strategies in teaching physics at upper secondary school in Sweden. They were interested in understanding the different strategies used in teaching physics, by focusing on teachers as a group, and reconstructing their collective habitus and cultural capital. They used Bourdieu's concept of life and habitus as tools for their analysis:

people, engraved in their bodies and minds, possess persistent system of dispositions shaped by their previous life, dispositions that allows them to act, think and orient themselves in the social world and which also controls their practices so that the same social world is recreated or changed. (Bourdieu and Passeron 1970, p. 19).

[människor, inristade i sina kroppar och sinnen besitter seglivade system av dispositioner formade av deras dittillsvarande liv, dispositioner som tillåter dem att handla,

tänka och orientera sig i den sociala världen och som dessutom styr deras praktiker så att samma sociala värld återskapas eller förändras].

Engström and Carlhed found three categories of teachers: (1) Promoters of Technology (39 % teachers), are people who work as agents for change in the teaching and learning of physics and emphasize the usefulness of physics for technological development; (2) Promoters for Citizenship (14 % teachers) are people who work as democratic and intellectual challengers of physics's status quo. They are characterized by engagement in society and political subjects and an interest in natural science and educational policy, emphasizing the concerns of physics for sustainable development. (3) Most physic teachers were categorized as *Managers of the Traditional* (46 % of teachers), the habitus of traditional teaching systems characterized by devotion to the field of natural science and a "sense of anxiety towards culture fields and resistance to social engagement. This habitus is characterized by closeness to natural science and unwillingness to contest and change. Hence, the teaching will include traditional methods to understand concepts and for succeeding with calculations using physics formulas" (p. 721).

Oral participation in the science classroom

According to Lindsay Mack (2012) oral participation or lack in the context of the classroom is a complex issue that not only involves cultural characteristics or linguistic abilities, but also instructional methodologies and social inclusion. She argues that sociocultural norms and values of teachers and students influence classroom discourse and that while student-centred rather than teacher-centred methodology encourages more participation, it provides no guarantees to learning as other factors may be involved. In a longitudinal study of the factors affecting implementation of argumentation, Anita Martin and Brian Hand (2009) found that the shift from teacher-centred to student-centred teaching methodology enhanced students' voices and some elements involved in science argumentation, such as dialog, design of investigations and making claims, and increased student-to-student dialogical interactions. However, Bryan Brown (2004) has examined the way students attending biology lessons are shaped by their peers and how those discourses require modification to scientific discourse. He explains that peers do not necessarily support each other's learning, and that some students are depressed and ridiculed by peers in class discussions because they used scientific language to explain scientific concepts. Brown reasons that students used their agency in whole class discussions deciding what kind of discourse to use by weighing the costs and benefits of using scientific language. Other students in his study engaged in scientific discourse and chose to ignore peers when talking in whole group discussions. Then, his peers may recognize this behaviour as socially unacceptable.

What a sociocultural approach to learning implies

A sociocultural approach of learning postulates that learning is embedded in social situations and interactions in which the actors appropriate cultural and discursive tools are mediated by language (Wertsch 1978). By participation in cultural activities individuals learn and develop through their shared use of cultural tools and social practice. Those social processes also involve processes of construction of identity (Lave and Wenger

1991). Thus, identity formation is not an isolated activity but it is an intrinsic aspect of learning.

Vygotsky's general genetic law of cultural development asserts that the most important factor in development of higher mental functioning in the individual is social character (Daniels 2008). According to this conception, learners first participate socially in the use of cultural tools and practices and then individually appropriate or "take up" the tools (Wertsch 1978). The law supports much of the subsequent and related theories that Vygotsky construed. The zone of proximal development (ZPD) is an example. It describes the difference between an individual's independent performance and his performance in cooperation with others (Daniels 2008). The ZPD emphasizes the distinction between inter, and intra mental functioning, i.e. performance that is externally mediated by interactions with others and performance that is internally mediated by interaction with the self (Daniels 2008). In this way, Vygotsky characterized development as a process that follows the learning process. His interest was in assessing the ways in which learners make progress under the process of instruction (Ibid.). According to Black (2003), the function of assessment in the Vygotskian perspective helps teachers and students to identify the zone of proximal development, directing students in new challenges. For teachers interpreting learner knowledge and abilities is expanded beyond surveillance of student solo acts, to learner contributions and responsiveness during activity with others with the assessor being crucial to the process. The quality of interactions in the process is crucial to helping learners move toward overcoming current difficulties (e.g., Poehner 2011).

The research design derives from real life narratives from students

A total of 15 students were interviewed and selected from the three classes involved in a previous ethnographic study. The interviews were face-to-face in the form of multiple group conversations. It involved a series of two interviews with every group with the aim of performing a gateway narrative inspired by Caroline Lundford Mears (2009). The multiple conversations allowed to me to maximize the communication and learning of the individuals who agreed to participate in this project. I used this method because it might provide an appropriate situation for quiet students to communicate their ideas and experiences with teachers' assessments, as members of the class, without being controlled or disturbed. The interviews in groups involving several individuals at a time have both conversation and interview character and are appropriate for analysing consistency and shared opinion in the group. It matches my aim in this study to induce the participants to reflect on their own experiences of learning and assessment of sciences, as symbolized by their conversations. According to Lundford Mears (2009), if students see their voices printed, it becomes their narrative, a validated experience, in particular because some students' voices may be under-represented in many contexts but not in the gateway narrative. The narrators' inspection of their voices is an essential factor in the reflection process.

Students from the first and third classes in biology, chemistry and physics were free to shape themselves into conversation groups of two, three or four individuals per time in each class. I called the students narrators and collaborators. Those conversations were performed at the end of class observations and were divided into two and sometimes three cycles of conversations. The first cycle began with informal conversations with groups of 2–4 students where I explained the aim of the interviews, these interviews were not

recorded, according to the recommendations of the gateway narrative by Lundford Mears. What was crucial in this initial conversation was to make the narrators feel comfortable and to gain their trust. In getting ready for the conversations, I also prepared the community by explaining what I was doing.

On one occasion during the first interview, a narrator in a group asked, at the same time they were looking around to ensure the teacher was not looking at her: “Are you sure that our chemistry teacher will not come to know what we are talking about?” “Yes I am”, I answered, and I explained again the ethical issues involved in this research. After a small pause, I told them that they could sue me by law if they felt cheated by me. Now they were happier and less anxious, then, it was easily to enter into conversation with this group of students. The ideal and full gateway narrative suggests a multitude of interviews but I did not have the opportunity to hold more than two and sometime three narrative interviews with each group because of time limits. The deep-interviews performed in group conversations with students were all audio-taped, transcribed. In the second round of interviews students were asked about their conception of what teachers assess for the purposes of grading, about students’ participation in the assessment process, specifically if there were some other kind of assessment such as peer- and self-assessment and why most of them were silent during teacher’s lecturing. Then I concluded the study by deep-interviews with the narrators when I had an opportunity to expand the issues I had introduced in the first and second conversation sequences.

Trying to understand the essence of what the students said

I am trying to understand students’ experiences through their participation in learning and assessment processes and their conceptions of teachers’ assessments by analysing what students said. According to the gateway narrative method by Lundford Mears (2009), the most important issue in the analysis is our connection with those persons whose experiences are being told in this study. The interviews speak directly to the reader to capture the intensity of the participants’ voices instead of being presented via codes for the purpose of preserving the discourses. It is a process of connecting with others where you learn from the experiences they told you which fits well with a sociocultural approach.

Thus when analysing gateway narratives, one must take care to treat that part of the students’ story seriously. I started by analysing what one of the groups said and tried to identify patterns in their narrative that matched with the other groups. Then I performed a cross-case analysis to shed light on what could be learned from the other groups of narrators: what did the data mean across all the narrators? Overlaying, coinciding or contradicting narratives were analysed and considered in the displayed findings. I went through the group narratives many times to notice patterns in the narrative, contradictions and differences among the diverse groups and to analyse and transcribe the data for everyone participating in the groups. In order to consider each narrative of specific interest to each of the research questions a separate document for each research question was developed and reviewed highlighting patterns related to every research question.

I was careful to avoid loss of nuances in the narrative in particular during deep interviews as I did not want to lose information but accurately reveal what are students’ feelings about both learning and assessment. Students’ feelings about the issue of assessment must be treated carefully and with consequence. For example I did not ask students if they had experience favouritism from teachers. However, some students in different groups without

being asked did bring up the issue of favouritism. The issue of teacher favouritism seems to be a sensitive issue for students. Specifically, teachers who practice favouritism irritated them.

Students' conceptions of teachers' assessments

The transcript answers the research question: What are students' conceptions of teachers' assessments?

Excerpt from group conversation with students on course I in biology

1. Interviewer: I am interested to know, what is your conception of teachers' assessment? [Jag är intresserad av att få veta Er uppfattning om lärarnas bedömning. Vilka uppfattningar har Ni om lärares bedömning?]
2. Patrik: What is assessment of something? [Vad är att bedöma för någonting?]
3. Interviewer: *Explains what assessment is.*
4. Patrik: Ah... Assessment... We don't really know. We have only got grades for the tests we have done and it is the grades you obtain in the test that count (...) what we have right because of our results on the written test and assignments. [Aha... bedömning, vi vet inte riktig. Vi har bara fått betyg på våra prov och det är betyg på provet vad som räknas (...) det är vad vi får, det är vad vi har rätt på grund av våra resultat på provet och redovisningar].
5. Interviewer: Humm... and what is the most important thing the teacher assesses in biology studies for the purpose of grading? [Humm... och vad är det viktigaste vad det gäller Era prestationer i biologi som lärarna bedömer hos Er i syfte att betygsätta Er i ämnet biologi?].
6. Kent: We don't know because we do not have any other kind of assessment, only our results in the tests. [Vi vet inte för att de ger inte oss en annan sort bedömning, bara våra resultat på provet].
7. Interviewer: Do you think that teachers assess other things in addition to the test and assignments? [Tycker Ni att lärarna bedömer andra saker förutom resultat på provet och redovisningar?].
8. Patrik: We don't know yet. [Vi vet inte än]
9. Jim: Yes, you must like them! ... Ha, ha, ha ha! [Ja, man ska gilla dem, he, he, he, he!].
10. Patrik: Yes, it could be teachers who have their favourites in the class, we don't know yet who they are. [Ja, det kan finnas lärare som favoriserar elever, det vet vi inte än vilka de är (...)].

Students do not know what the word assessment means: *What is assessment of something?* (2). The concept of assessment by a specialist involving planned observation and gathering of information about students' performances is unknown by biology student Patrik. We have explained in the preliminary conversations how assessment is conceived in the research and in the curriculum but the culture of grading seems to be rooted. Grading is the term most used by students. The only kind of assessment students have got is the result of their written paper-and-pencil tests: *We have only got grades for the tests we have done and it is the grades you obtain in the test that count* (4). Students have not experienced any

kind of assessment other than the written tests: *we do not have any other kind of assessment, only our result in the tests*. (5). Students have some conception of assessments being associated with favouritism: *Yes, you must like them!... Ha, ha, ha ha!* (10). Other in the group agreed, *Yes, it could be teachers who have their favourites in the class, we don't know yet who they are* (11).

What teachers actually assess for the purpose of grading

The following transcript answers the research questions: What is particularly important for the teachers to assess above all for the purposes of grading? Do you think that teachers assess other things in addition to the test and other assignments?

Group conversation with students in the upper secondary school course I chemistry

1. Interviewer: I would like to know what are the most important things the teachers assess for the purpose of grading? [Jag skulle vilja veta vad är det viktigaste lärarna bedömer hos Er i syfte att betygsätta Er]
2. Rose: The written test. [Provet]
3. Elena: Yeah, the written test, assignments and how much we participate in the lesson. [Just det, provet, redovisningar och hur pass aktiva vi är under lektionerna]
4. Interviewer: Do you think that teachers assess other things in addition to the written test and assignments? [Tycker Ni att lärarna bedömer andra saker förutom resultat på provet och redovisningar?].
5. Rose: Some teachers grade behaviour but I do not think that they should. [Vissa lärare betygsätter uppförande men det tycker jag inte att lärarna ska göra det],
6. Elena: Yes, some teachers do but not all, but we don't know how they do. [Ja, vissa lärare gör det men inte alla, men vi vet inte hur de gör].
7. Rose: Yes, I sometimes feel that the teachers favour certain students. [Ja, jag upplever ibland att lärarna favoriserar vissa elever]. Elena: Yes, they do! [Ja, det gör de!].

The written examination is the most important assessment criterion teachers used: *The written test* (2). The other student agreed but added: *and how much we participate in the lesson* (3). But this student could not explain how this assessment is performed by the teacher because she did not know. Student Rose has experienced teacher assessment other than assessment of learning: *Some teachers grade behaviour but I do not think that they should* (5). These two students were in agreement on some conceptions of teachers' assessments existing in the first group conversation, namely, about the favouritism of the teachers for some students: *I sometimes feel that the teachers favour certain students* (7, 8). Another common conception in the first group conversation was the expression of inexperience in differentiating the assessment from grading process.

Students fear being labelled and group pressure

In the following transcript students on the course I in chemistry give answers to the research question: Why are most of the students silent during the teachers' lecturing?

Students' silence during the teachers' lecturing—group conversation with students on the course I in chemistry

1. Interviewer: During the lessons students do not often participate orally or in discussions. Students are pretty silent. What do you think can be the reason for this? [Under genomgångarna oftast deltar ni inte i diskussioner utan Ni är ganska tysta, vad kan det bero på tror Ni?].
2. Angie: Yes, you don't want to participate in discussions, you prefer sitting back, you don't want to be seen. [Ja, man vill inte delta i diskussionerna, man vill gärna sitta längst bak och inte säga nåt, man vill inte synas].
3. Interviewer: Why is this situation so, do you think? [Varför är det så, tror Ni?].
4. Lisse: Yes, it is hard to hold up one's hand because of group pressure. Even if you know the answer to the teacher's question, you don't want to talk. You want your classmates to believe you are cool, otherwise they can say you are a jerk, because you ask questions. [Ja, vi tycker att det är jobbigt att räcka upp handen och det är på grund av grupptycket. Även om man kan svaret på frågan läraren ställer, så vill man inte prata. Man vill visa sig i klass-rummet att man är cool, annars kan de andra säga att man är en tönt som frågar sant].
5. Interviewer: Hummmm.
6. Angie: Yes, you do not dare to talk even if you can answer a question. [Ja, man vågar inte prata även om man kan svaret på en fråga].

This group conversation with students shows students do not dare to talk during the lecture because of group pressure. Students claim that it is not cool to talk in the classroom, and they fear being labelled by peers: *You want your classmates to believe you are cool, otherwise they can say you are a jerk, because you ask questions* (4). Group pressure is obvious as students prefer sitting back: *you don't want to be seen* (2). Students claim that even when they want to participate in the discourse, they do not do it because of fear: *you do not dare talk even if you can answer a question* (6).

Students' conceptions on learning sciences

The following is a group conversation with three physics students on the science programme on course III in physics. They answer the research question: Why are most of the students silent during the teachers' lecturing?

Group conversation with three physics students on the science programme on course III in physics

1. Interviewer: During the lessons students do not participate in discussions. Students are pretty quiet. What can be the reason for this, do you think? [Under genomgångarna oftast deltar ni inte i diskussionerna, Ni elever är ganska tysta, vad kan det bero på tycker Ni?]
2. Cecilia: It is very easy to get the label "stupid" if you ask questions. You can be afraid to ask stupid questions. [Man får lätt en viss "dum" stämpel om man frågar, man kan vara rädd att ställa en dum fråga].
3. Interviewer: Hummmm

4. Eva: Yes, I agree, you are afraid to ask stupid questions. You do not do it because you are afraid. [Ja, det stämmer, man är rädd att ställa dumma frågor...man gör inte det av rädsla].
5. Susanna: Yes, there are many peers who want to hold up their hands and they do not do it because of fear. [Ja, det är många som hade velat räcka upp handen och som inte gör det av rädsla].
6. Cecilia: We also grew up with just being ‘spoon-fed’ sitting still and listening to teachers’ lecturing. [Vi är också uppväxta med att vi blir bara matade, sitta stilla och lyssna på genomgångarna].
7. Eva: It also depends on how eager to learn you are. [Det bero också på hur man vetgirig man är].
8. Susanna: Yes, but it depends also on how much you manage to understand, because if you really want to understand then you ask, isn’t that right? [Ja, men det bero också på hur mycket man orkar förstå, för att, vill man verkligen förstå så frågar man, eller hur?].

Just as the students in the course I, students in the course III of physics claim they do not participate in the discourse of science because they are afraid to be labelled by peers: *It is very easy to get the label “stupid” if you ask questions, you can be afraid to ask stupid questions* (2). They are also afraid to participate because of fear: *Yes, there are many peers who want to hold up their hands and they do not do it because of fear* (5). Students consider that their silence is to do with the culture of their classroom. They are familiarized and cultured with the pedagogic modality to be silent and to be fed by teachers: *We also grew up with just being ‘spoon-fed’, sitting still and listening to teachers’ lecturing* (6). It seems that student silence during the teachers’ lecture was because they were afraid to ask the teacher questions and to participate in discussions. These students’ ideas on learning sciences as if being fed by teachers may be interpreted as a link between their silence and their way of perceiving learning in science.

This finding is important to remember when discussing how teachers can use students’ participation in classroom discussions. According to Dylan Wiliam (2000) it is student participation and the activities that would be likely to generate dialog which are important and have a value in themselves and not the correctness of the students’ questions and answers. However, this finding shows that students are afraid to participate in discussions and this may suggest that the environment for discussions is not a productive learning environment. This depends in turn on how prepared teachers are or not to support students’ learning and how teachers need to approach students’ learning (see Brookhart 2011).

Student’s conceptions about learning sciences—silence and traditional science learning

The following transcript answers the research question: Why are most of the students silent during the teachers’ lecturing?

Group conversation with physics students in the course III

Interviewer: During the lessons often students do not participate orally or in discussions while the teacher is lecturing. Students are pretty silent. What can be the reason for this do you think? [Under genomgångarna oftast deltar ni inte i diskussioner utan Ni är ganska tysta, vad kan det bero på tror Ni?].

1. Robert: It depends if you understand everything the teacher is saying, then there is nothing to ask. [Det kan bero på man förstår allt läraren säger, finns ingenting att fråga]
2. John: It depends on the subject; if the subject requires that you must have opinions, for example social studies and history, then you must talk. It is different in physics, in the experiment work it may be possible to ask the teacher but not during the lecturing. [Det bero på ämnet, om det gäller att ha åsikter det är mening att man ska ha åsikter, till exempel i samhällskunskap och historia, men det är annorlunda i fysik, på labbet kanske kan man diskutera men inte på lektionerna].
3. Charles: Yes, I agree. [Ja, jag håller med].

The transcript discloses students' conceptions on learning sciences and may connect student silence with students' perceptions of how to learn sciences: *It depend on the subject, if the subject requires that you must have opinions, for example social studies and history, then you must talk, it is different in physics, [...]* (3). This is typical reasoning by students who are cultured to traditional science education that was found in earlier research to be the practice of teachers belonging to the habitus *Manager of the Traditional*, according to Engström and Carlhed (2014). Like their science teachers, students state that they do not consider it necessary either to discuss different cultural issues in science; the field of natural science does not include cultural discussions or discussions on social engagement.

Students' experience of peer assessment

The following transcript shows the conception students have about assessment at the *impact* level. We asked about students' experiences of some characteristics of formative assessment, namely peer-assessment. The transcript answers the research question: What is the students' experience of peer-assessment?

Group conversation with students on course I in biology

1. Interviewer: I would like to know what is your experience of peer-assessment? [Jag skulle vilja veta vad vet Ni om kamrat bedömning?]
2. Marcus: What is assessment of something? [Vad är bedömning för någonting?]
3. Interviewer: *Gives an explanation about assessment and peer-assessment.*
4. Marcus: Ok [...] to assess our peers can be difficult. Perhaps, if you assess your nearest peers, because the others, you don't know so much about them. [Ok... att bedöma vår klasskamrat kan vara svårt, kanske om man bedömer sin närmaste kompis, för att de andra, man vet inte så mycket om de].
5. Johannes: Yes, to listen to other voices about our own achievement could be good for us, but we don't have such a system in our school. It is only the teacher who assesses

our outcomes. [Ja, att lyssna andra röster om ditt eget arbete kan vara bra för oss, men vi har inte sådan i skolan. Det är bara läraren som bedömer vårt arbete].

Students seem not to have an idea about what assessment is, even though I explained what assessment is during the first interview conversation. As a result, I used their own concepts of assessment, i.e. scores: *What is assessment of something?* (2). After an explanation about what peer assessment is I found that students took interest: *Yes, to listen to other voices about our own achievement could be good for us* (5). However, as students have not had experience of peer-assessment they do not have clear ideas about how it works and how to implement it. *Ok [...] to assess our peers can be difficult. Perhaps if you assess your nearest peers, because the others, you don't know so much about them* (4). This finding may be interpreted as a lack of training for students to practise peer-assessment.

Students' experiences of self-assessment

The transcript answers the research question: What is students' experience of self-assessment?

Group conversation with three physics students in the course III

1. Interviewer: I would like to know what is your experience of self-assessment? [Jag skulle vilja veta vika kunskaper har Ni om själv bedömning?]
2. Maria: What is self-assessment? I have never heard of it. [Jag har aldrig hört om det, vad är det?].
3. Interviewer: *Explained the term.*
4. Maria: It is difficult. It depends on how accurate you are in your self-assessment... you cannot say to the teacher: "I want to get a *pass with distinction*". [Det är svårt, det bero på hur pass mycket noggrant du är i din egen bedömning...Du kan inte säga till läraren: "Jag vill ha MVG"].
5. Angelica: Yes, you must have understanding about how good you are if you are to assess yourself. [Ja, du måste ha insikt om hur bra du är om du ska kunna bedöma dig själv].
6. Maria: I don't really know, because you can be self-righteous, but perhaps you are not. [Jag vet inte riktigt, för att du kan vara självgod, men kanske du inte är det].
7. Susanne: Yes, but it would have been good for us to participate in the assessment process, but we do not. [Ja men det skulle ha varit bra för oss delta i bedömnings processen men vi gör inte det].
8. Interviewer: But it is something you might be able to develop into. [Men det är nåt man kanske skulle kunna utvecklas till?].
9. Maria: Yes, it's like a maturation phase. [Ja, det är det är som en mognadsfas].
10. Angelica: Yes, but to discuss with the teacher about ones grades... there are many peers which are good at talking to teachers about their grades, they are good at marketing themselves: "I have done this very well," they may say, but others may be more withdrawn and shy and in that case the person can become... I don't know. [Ja, men detta med att diskutera med läraren om sina betyg... det är många som är bra för att prata med lärarna om sina betyg, att de är bra att marknadsföra

sig själv: ”jag har gjort detta jätte bra”, men andra kanske är mer tillbakadragen och blyg i så fall blir den person... jag vet inte].

11. Susanne: Yes, I think it is important to discuss grades with the teacher. [Ja, jag tycker att det är viktigt att diskutera betyget med läraren].

What student Cecilia expresses represents a picture of some misconceptions of assessment the students have. She confuses assessment at the judgement level (grades) with assessment at the impact level (formative assessment) (Newton 2007), and she said: *you cannot say to the teacher: “I want to get a pass with distinction”* (4). Her assertion applies first and foremost to summative assessment and not to formative assessment at the impact level. The culturization on self-assessment is an issue requiring planning, communication with the students and collaboration between teachers and students (Harrison 2011), so what the students claim may be interpreted as a lack of training in self-assessment. The disorientation of students probably leads them to talk about *grades* when you have asked about *assessment*. But these students are reflective and capture the direct meaning of formative assessment: *Yes, you must have understanding about how good you are if you have to assess yourself* (5); and at the same time are hesitant about something they are not accustomed to in practice: *I don’t really know, because you can be self-righteous, but perhaps you are not* (6). Instruction on formative assessment for both teachers and students may give most of the students opportunities to participate in their assessment and prevent inequalities in the assessment process, because some of them are good at marketing themselves: *Yes, but to discuss with the teacher about their grades... there are many peers who are good at talking to teachers about their grades, they are good at marketing themselves: “I have done this very well,” they can say, but others may be more withdrawn and shy and in that case the person can become... I don’t know* (10).

Students take the implementation of formative modalities of assessment positively: *Yes, it would have been good for us to participate in the assessment process, but we do not* (7). Some opponents of self-assessment claim that students probably cannot assess their own performance in an objective way, but according to Wiliam (2000), their assertion applies first and foremost to summative assessment. He claims that precision in formative assessment is an issue of secondary significance and that what actually matters in formative assessment is whether self-assessment can enhance learning.

What I have learned and how I have reflected

The students were silent during teachers' lectures because they were afraid of group pressure, feared disappointing their peers and afraid to get a silly stamp. Other motives of their silence is that they considered that in science one does not need to discuss or argue content or knowledge and that they are accustomed to sitting in silence listening and being fed by the teachers. The finding may suggest that students are not trained to talk sciences and the opportunities to challenge, propose or dispute natural science issues are limited by teachers' lead lessons. The silent teacher-centred milieu with rather passive students is not in accordance with advocating democratic or scientific working methods, which are in general based on openness, critique, respect for multiple stand points and views and assume clear, transparent argumentation from different parties, aiming at a mutual understanding. In the case of those students experiences the dominant teacher habitus is confirmed, *Manager of the Traditional*, is predominant matching the results of this study with the results of Engström and Carlhed (2014).

Since the students have limited, if any, knowledge of assessment besides written tests this may suggest some degree of distrust in the fairness of grading in science classes. Partially because students do not know what are the evidences of learning collected in order to obtain their grades. Moreover the written test has limitation about the possibilities to assess different qualities of knowledge, as for example critical thinking, argumentation or analytic skills, important qualities on methods of scientific work when learning sciences. Moreover, the collection of evidences, as for example analytical skills on those kinds of milieu is prevented by the limited opportunities to talk science in classroom.

Based on the student voices, an inevitable conclusion from this research is that there are major discrepancies between what the Swedish national curriculum establishes about assessment and assessment approaches in science education. Although the sample is little, the school is not an exceptional setting or deviating in any way, the school is an ordinary upper secondary school and representative of corresponding classes across the country. There is no research that suggests that the founding may not be consistent with how appears to be or other that have report that the findings are different or abnormal. Possible causes for the discrepancies found remain to be discovered and may well be linked to areas of research besides pedagogy. Clearly, forces behind the discrepancies have a direct impact on pedagogy, assessment and grading, as performed in Swedish science classes.

What assessment specialists recommend is that assessment is a continuous process wherein teachers sit with students and they build together cooperative agreements during the learning process (e.g., Lundahl 2011). According to Christian Lundahl (2011) teacher's role is crucial if assessments shall promote or inhibit learning. Susanne Brookhart (2011) argues for the use of formative assessment with purposes to support learning and pointed out that teachers need create formative assessment environments but that also teachers need support and instruction to change outdated assessment methodologies. Several studies in the field of assessment show that formative interaction influences the internalization of knowledge in many aspects (e.g. Lundahl 2011) and these, in turn, raise the standards of achievement. Unfortunately, research in this field has yet to be conducted.

References

- Black, P. (2003). The importance of everyday assessment. In J. M. Atkin & J. Coffey (Eds.), *Everyday assessment in the science classroom: Science educators' essay collection* (pp. 1–12). Arlington, VA: National Science Teachers Association.
- Black, P., & Wiliam, D. (2003). "In praise of educational research": Formative assessment. *British Educational Research Journal*, 29(5), 623–638. doi:[10.1080/0141192032000133721](https://doi.org/10.1080/0141192032000133721).
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5–31. doi:[10.1007/s11092-008-9068-5](https://doi.org/10.1007/s11092-008-9068-5).
- Blatchford, P., Baines, E., Basset, P., Chowne, A., & Rubie-Davies, C. (2006). The effect of a new approach to group work on pupil-pupil and teacher-pupil interactions. *Journal of Educational Psychology*, 98(4), 750–765. doi:[10.1037/0022-663.98.4.750](https://doi.org/10.1037/0022-663.98.4.750).
- Bourdieu, P., & Passeron, J. C. (1970). *Reproduktionen*. Sweden: Moderna klassiker. Arkiv förlag. Översättning Gunnar Sandin. Lund.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.
- Brown, A. B. (2004). Discursive identity: Assimilation into the culture of science and its implication for minority students. *Journal of Research in Science Teaching*, 41(8), 810–834. doi:[10.1002/tea.20228](https://doi.org/10.1002/tea.20228).
- Cizek, G. J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice*, 92(2), 16–23. doi:[10.1111/j.1745-3992.2000.tb00026.x](https://doi.org/10.1111/j.1745-3992.2000.tb00026.x).
- Daniels, H. (2008). *Vygotsky and research*. New York: Routledge.

- Engström, S., & Carlhed, C. (2014). Different habitus: different strategies in teaching physics? Relationships between teachers' social, economic and cultural capital and strategies in teaching physics in upper secondary school. *Cultural Studies of Science Education*, 9, 699–728. doi:[10.1007/s11422-013-9538-z](https://doi.org/10.1007/s11422-013-9538-z).
- Forsberg, E. & Lindberg, V. (2010) Svensk forskning om bedömning—En kartläggning. Vetenskapsrådet. *Vetenskapsrådet rapportserie*, 2:2010. Stockholm, Sweden: The Swedish research on assessment—a mapping. *The Swedish Council of Science Research* 2:2010, 1–135.
- Gómez, M., & Jakobsson, A. (2014). Everyday classroom assessment practices in science classrooms in Sweden. *Cultural Studies of Science Education*,. doi:[10.1007/s11422-014-9595-y](https://doi.org/10.1007/s11422-014-9595-y).
- Harrison, C. (2011). Peer- and self-assessment. In Sanna Järvelä (Ed.), *Social and emotional aspects of learning* (pp. 169–173). Oxford, UK: Academic Press.
- Haug, P. (2003) *Evaluering av Reform 97. Slutrapport frå styret for Program for evaluering av Reform 97*. Oslo, Norway: Norwegian Research Council.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lavonen, J., Angell, C., Ellen, R., Bymen, K., Henriksen, E. K., & Koponen, I. T. (2007). Social interaction in upper secondary physics classrooms in Finland and Norway: A survey of students' expectations. *Scandinavian Journal of Educational Research*, 51(1), 81–101. doi:[10.1080/00313830601079082](https://doi.org/10.1080/00313830601079082).
- Lundahl, C. (2011). *Bedömning för lärande [Assessment for Learning]*. Stockholm: Nordstedt. Sweden.
- Lundford Mears, C. (2009). *Interviewing for education and social science research. The gateway approach*. London: Palgrave Macmillan.
- Lysne, A. (2006). Assessment theory and practice of students' outcomes in the Nordic countries. *Scandinavian Journal of Educational Research*, 50(3), 327–359. doi:[10.1080/00313830600743365](https://doi.org/10.1080/00313830600743365).
- Mack, L. (2012). Does every student have a voice? Critical action research on equitable classroom participation practices. *Language Teaching Research*, 16(3), 417–434. doi:[10.1177/1362168812436922](https://doi.org/10.1177/1362168812436922).
- Martin, A. M., & Hand, Brian. (2009). Factors affecting the implementation of Argument in the elementary science classroom. A longitudinal case study. *Research in Science Education*, 39, 17–38. doi:[10.1007/s11165-007-9072-7](https://doi.org/10.1007/s11165-007-9072-7).
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 2, 149–170. doi:[10.1080/09695940701478321](https://doi.org/10.1080/09695940701478321).
- Norrild, P., Angell, C., Bang, H., Larsen, C., Paulsen, A., & Stubgaard, S. (2001). *Fysik i skolen skolen i fysik. Evaluering af fysik i det almene gymnasium*. Copenhagen: Danmarks evalueringsinstitut.
- Nyström, E. (2007) *Talking and taking positions. An encounter between action research and the gendered and racialised discourses of school science*. Doctoral dissertation in Pedagogy, Nr 16. [Doktorsavhandlingar i Pedagogiskt arbete Nr 16]. Umeå: Umeå University, Sweden.
- Poehner, M. E. (2011). Dynamic assessment: Fairness through the prism of mediation. *Assessment in Education: Principles, Policy & Practice*, 18(2), 99–112. doi:[10.1080/0969594X.2011.567090](https://doi.org/10.1080/0969594X.2011.567090).
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5, 77–84. doi:[10.1080/0969595980050104](https://doi.org/10.1080/0969595980050104).
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). Cooperative learning and achievement. In W. M. Reynolds & G. J. Miller (Eds.), *Handbook of psychology, vol 7: Educational psychology* (pp. 177–198). Hoboken, NJ: Wiley. doi:[10.1002/0471264385.wei0709](https://doi.org/10.1002/0471264385.wei0709).
- Stipulation 1980:64 [Förordning 1980:64]. Lgr 1980. *Swedish Curriculum Compulsory School (1980)*. Stockholm: Swedish Ministry of Education.
- Swedish Ministry of Education. (1994). *Swedish National curriculum*. Stockholm: Swedish Ministry of Education.
- Swedish Ministry of Education. (2013). *Swedish National curriculum*. Stockholm: Swedish Ministry of Education.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education*, 14(3), 281–294. doi:[10.1080/09695940701591867](https://doi.org/10.1080/09695940701591867).
- Wertsch, J. V. (1978). *Voices of the main: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.
- Wiliam, D. (2000). Formative assessment in mathematics. Part 3: The Learner's role. *Equals Mathematics and Special Educational Needs*, 6(1), 19–22.
- Wiliam, D., & Black, P. (1996). Meaning and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537–549. doi:[10.1080/0141192960220502](https://doi.org/10.1080/0141192960220502).

María del Carmen Gómez is a PhD student at the Faculty of Social Sciences at Lund University, in Sweden. Her research is based on exploration of teachers' assessment practices in science education in Sweden and her work is grounded in sociocultural theories of learning and assessment. She believes strongly that natural sciences should be available for all students and that assessment should be used to enhance learning.

DOKTORSAVHANDLINGAR I PEDAGOGIK, LUNDS UNIVERSITET

Lennung, Sven-Åke. *Meta-learning, laboratory training, and individually difference change*. Malmö: PA-Council. 1974.

Nihlén, Kerstin. *Samspelet mellan skola och hem. Former för kontakt och samarbete mellan lärare, föräldrar och elever i grundskolan*. Stockholm: Bonniers. 1976.

Bergsten-Brucefors, Agneta. *Intelligence Patterns and their relation to social background. A longitudinal study*. Lund: Gleerup. 1976.

Bratt, Ingar. *Engelskundervisningens framväxt i Sverige. Tiden före 1850*. Uppsala: Årsböcker i svensk undervisningshistoria, 139. 1977.

Drakenberg, Margareth. *Förändringar i prestationsprocessen. En longitudinell studie av skolbarn från 10-13 år*. Lund: Gleerup. 1978.

Eriksson, Mona. *Utbildningspraktik. Om utbildning med handledda klient- och patientkontakter på högskolenivå*. Lund: Gleerup. 1978.

Flinck, Rune. *Correspondence education combined with systematic telephone tutoring*. Lund: Gleerup. 1978.

Norberg, Astrid. *Uppfostran till underkastelse. En analys av normer för föräldrabarnrelationer i religiös litteratur om barnuppfostran i Sverige 1750-1809*. Lund: Gleerup. 1978.

Carlsson, Bertil. *Två utbildningsprojekt. En redovisning av arbetet att profilera verksamheten vid en nystartad svensk högskola*. Lund: Gleerup. 1979. (Disp 1980)

Bååth, John A. *Postal two-way communication in correspondence education*. Lund: Gleerup. 1980. (Disp 1981)

Torper, Ulf. *Tidsramar, tidsanvändning och kunskapsutveckling i den svenska grundskolan*. Lund: Gleerup. 1982.

Eppler, Mats och Nelander, Bo. *Kartläggning och omformning av den psykosociala arbetsmiljön på ett sjukhus*. Malmö: Liber. 1984. (Disp 1985)

Ritchey, Elisabet. *Pedagogisk teori och vetenskapspolitik. Tre historiska strömningar*. Malmö: Liber. 1985.

Selander, Sven-Åke. *Livslångt lärande i den svenska kyrkoförsamlingen. Fleninge 1820-1890*. Uppsala: Erene. 1986.

Nilsson, Urban. *Pensionärer och studieverksamhet*. Lund: Pedagogiska institutionen. 1986.

Lundbergh, Beate. *Kom ihåg att du är underlägsen! Pedagogik för borgarflickor i 1880-talets Sverige*. Lund: Studentlitteratur. 1986.

Nilsson, Ingrid. *En spjutspets mot framtiden. En analys av de svenska enhets- och grundskoleformerna i utländsk vetenskaplig litteratur 1950-1980*. Lund: Pedagogiska institutionen. 1987.

Ekwall, Sven. *Tidig småskolläraryr utbildning. En studie med särskilt avseende på Malmölänet 1865-1884*. Uppsala: Årsböcker i svensk undervisningshistoria, 161. 1987.

Härnsten, Gunilla. *Ett decennium av SOL i Sverige. Socialistisk lärarorganisering på 1970-talet i Sverige, med jämförande exempel från några andra länder*. Lund: Pedagogiska institutionen. 1988.

Lundahl, Lisbeth. *I moralens, produktionens och det sunda förnufts namn. Det svenska högerpartiets skolpolitik 1904 - 1962*. Lund: Pedagogiska institutionen. 1989.

Lundberg, Per. *Utbildning och träning för interkulturell kommunikativ kompetens*. Lund: Pedagogiska institutionen. 1991.

Bauer Alfredsson, Betty. *Förändringsarbetet i äldrevården. Utvärdering av personalutbildning inom institutionsbunden och alternativ äldrevård*. Lund: Studentlitteratur. 1991.

Holm, Kerstin I M. *Företagsutbildning. Tillvaratagande av de mänskliga resurserna?* Lund: Pedagogiska institutionen. 1991.

Erlöv, Iris och Pettersson, Kerstin. *Från kall till personlighet? Sjuksköterskans utbildning och arbete under ett sekel*. Lund: Pedagogiska institutionen. 1992.

Gerrevall, Per. *Högskolestuderandes erfarenheter av självständigt arbete*. Lund: Pedagogiska institutionen. 1992.

Deijenberg, Margot. *Pedagogik och människosyn: Om barnlitteratur och bildningssträvanden i 1860-talets Sverige*. Lund: Pedagogiska institutionen. 1992.

Liljedahl, Kerstin. *Handikapp och Omvärld - 100 års pedagogik för ett livslångt lärande*. Lund: Pedagogiska institutionen. 1993.

Kreitzberg, Peeter. *The Legitimation of Educational Aims: Paradigms and Metaphors*. Lund: Pedagogiska institutionen. 1993.

Magano Amukugo, Elizabeth. *Education and Politics in Namibia. Past Trends and Future Prospects*. New Namibia Books. 1993.

Johansson, Bertil & Johansson, Britt. *Att styra eller inte styra. En kritisk granskning av den svenska grundskolans förändrade styrning med inriktning mot planeringsdokument*. Lund: Pedagogiska institutionen. 1994.

Johansson, Margareta. *Det är ju ändå vi som ska ta över världen. Om grundskole- ungdom och den nya fredsfostran i 1980-talets Sverige*. Lund: Pedagogiska institutionen. 1994.

Bendz, Mona. *Kunskap i praktik. Hur sjuksköterskestuderande uppfattar kliniska vårdsituationer och sitt eget handlande i dem samt hur deras uppfattningar förändras under och efter utbildningen*. Lund: Pedagogiska institutionen. 1995.

Holmqvist, Mona. *Autism. Uppfostran, undervisning och förståelse för personer med extremt atomistisk omvärldsuppfattning*. Lund: Pedagogiska institutionen. 1995.

Ohlin, Christer. *Special education in Oman. Primary school teachers' perception of children with learning disabilities*. Stockholm: Almqvist & Wiksell International. 1995.

Härdig, Jan. *Att utbilda till arbetare. En studie av gymnasieskolans bygg- och anläggningstekniska linje och yrkeslärare*. Lund: Pedagogiska institutionen. 1995.

Lindberg-Sand, Åsa. *Spindeln i klistret. Den kliniska praktiken betydelse för utveckling av yrkeskompetens som sjuksköterska*. Lund: Pedagogiska institutionen. 1996.

Chaib, Christina. *Ungdomsteater och personlig utveckling. En pedagogisk analys av ungdomars teaterskapande*. Lund: Pedagogiska institutionen. 1996.

Svederberg, Eva. *Tänkande bakom val och användning av livsmedel. Faktorer som medverkar till eller utgör hinder för förändring av matvanor i hälsofrämjande riktning*. Lund: Lund University Press. 1997.

Kindeberg, Tina. *Undervisningens möjligheter att förändra elevernas tänkande inom området aids och sex*. Lund: Lund University Press. 1997.

Vergara, Maria. *Silence, Order, Obedience and Discipline. The educational discourse of the Argentinean Military Regime (1976-1983)*. Lund: Lund University Press. 1997.

Ström, Peter. *Förändringsarbete och lärande. Om utveckling av förändringspraktik bland vårdbiträden i hemtjänsten*. Lund: Lund University Press. 1997.

Ahlner Malmström, Elisabet. *En analys av sexåringars bildspråk - Bilder av skolan*. Lund: Lund University Press. 1998.

Karlsson, Lars. *Leka, lära, öva, verkskapa. Elsa Köhler en österrikisk aktivitetspedagog i Sverige*. Lund: Lund University Press. 1998.

Fritzén, Lena. *Den pedagogiska praktiken janusansikte. Om det kommunikativa handlandets didaktiska villkor och konsekvenser*. Lund: Lund University Press. 1998.

Linder, Karin. *Perspektiv i sjuksköterskeutbildningen. Hur en grupp studerandes uppfattning av sjuksköterskans yrke förändras under tre år av utbildning*. Lund: Lund University Press. 1999.

Anderberg, Elsie. *The relation between language and thought revealed in reflecting upon words used to express the conception of a problem*. Lund: Lund University Press. 1999.

Szybek, Piotr. *Staging science. Some aspects of the production and distribution of science knowledge*. Lund: Pedagogiska institutionen. 1999.

Ekstrand, Britten. *Småbarnsskolan. Vad hände och varför? En sekellång historia studerad med fokus på förändring av pedagogisk verksamhet från 1833 och framåt*. Lund: Pedagogiska institutionen. 2000.

Rosenlind, Märta. *Vuxnas tänkande om och förhållningssätt till munhälsa*. Lund: Pedagogiska institutionen. 2000.

Lindberg, Berit. *Kvinnor - vakna, våga! En studie kring pedagogen och samhällsvisionären Honorine Hermelin Grönbech*. Lund: Pedagogiska institutionen. 2000.

Hansson, Birgit. *Förutsättningar för gymnasieelevers kunskapsbildning och för undervisning inom miljöområdet*. Lund: Pedagogiska institutionen. 2000.

Frisdal, Bodil. Lyhördhet. *Studenters uppfattningar av lyhördhet i omvårdnad och utbildning*. Lund: Pedagogiska institutionen. 2001.

Plöjel Westmoreland, Elisabet. *Tvärkulturell fostran. Svensk-italienska ungdomars erfarenhet av att växa upp med två kulturer*. Lund: Pedagogiska institutionen. 2001.

Hellström, Esbjörn. *Reformpedagogik i möte med den statskommunala skolan i Barnets århundrade*. Lund: Pedagogiska institutionen. 2002.

Persson, Ingvar. *Skolledare i grundskolan. En fallstudie av biträdande rektorers möte med skolledning*. Lund: Pedagogiska institutionen. 2002.

Chekol, Ing-Marie. *Handledning som undervisningsform i sjuksköterskeprogrammets praktik – en beskrivning av variation i innebörd*. Lund: Pedagogiska institutionen. 2003.

Berggren, AnneCharlotte. *“Hade jag inte läst de där artiklarna... “ Barnmorskors förhållningssätt till användning av forskningsresultat inom omvårdnad*. Lund: Pedagogiska institutionen. 2003.

Bager-Charleson, Sofie. *The Parent's School. Narrative Research about Parental Involvement in School*. Lund: Pedagogiska institutionen. 2003.

Wihlborg, Monne. *A Pedagogical Stance on Internationalising Education: An empirical study of Swedish nurse education from the perspectives of students and teachers*. Lund: Pedagogiska institutionen. 2005.

Båtshake, Helene. *Lena Lär Lyda. Fostran och disciplinering i svensk skola 1947-1956*. Lund: Pedagogiska institutionen. 2006.

Nilsson, Ingrid. *Grundskollärares tankar om kompetensutveckling*. Lund: Pedagogiska institutionen. 2006.

Bäckström, Lena. *Barn i bon. En intervjustudie med 11-åringar om familj, fostran och dagligt liv*. Lund: Pedagogiska institutionen. 2006.

Jedemark, Marie. *Lärarutbildningens olika undervisningspraktiker. En studie av lärarutbildares olika sätt att praktisera sitt professionella uppdrag*. Lund: Pedagogiska institutionen. 2007.

Brodin, Eva. *Critical Thinking in Scholarship: Meanings, Conditions and Development*. Lund: Pedagogiska institutionen. 2007.

Lorentz, Hans. *Talet om det mångkulturella i skolan och samhället: en analys av diskuser om det mångkulturella inom utbildning och politik åren 1973-2006*. Lund: Pedagogiska institutionen. 2007.

Porath Sjö, Elisabeth. *Konfirmandernas bildningsresa. Ungdomars berättelser om sitt deltagande i konfirmandundervisningen*. Lund: Pedagogiska institutionen. 2008.

Karlsson, Jan. *Learning in Collaboration. Academics' experiences in collaborative partnerships with practitioners and researchers*. Lund: Pedagogiska institutionen. 2008.

Löfgren Martinsson, Maria. *Högre utbildning och arbete med personal- och arbetslivsfrågor – om professionalisering och utveckling av anställningsbarhet*. Lund: Pedagogiska institutionen. 2008.

Johansson, Magnus. *Att bära världen på sina axlar. Miljövetares uppfattningar av och förhållningssätt till miljöproblem och sin profession*. Lund: Pedagogiska institutionen. 2008.

Dash, Ingrid. *Flexibility in knowing school mathematics in the contexts of a Swedish and an Indian school class*. Lund: Pedagogiska institutionen. 2009.

Olsson, Anders. *Konstruktionen av gymnasieskolans turismämne*. Lund: Pedagogiska institutionen. 2009

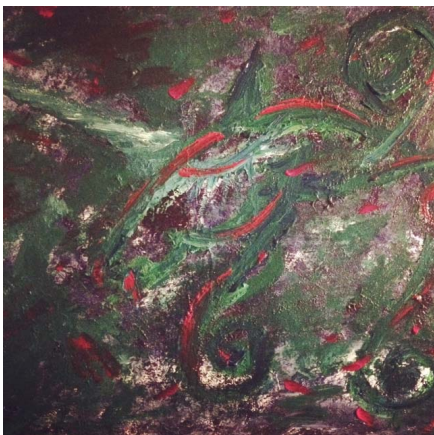
Åkerblom, Annika. *Children making sense of physical phenomena*. Lund: Avdelningen för pedagogik. Sociologiska institutionen. 2011.

Doumas, Kyriaki. *Students' experiences and perceptions of good teaching practice*. Lund: Avdelningen för pedagogik. Sociologiska institutionen. 2011.

Johansson, Lotta. *Tillblivelsens pedagogik: Om att utmana det förgivettagna. En postkvalitativ studie av det ännu-icke-seddans pedagogiska möjligheter*. Lund: Avdelningen för pedagogik. Sociologiska institutionen. 2015.

Ceder, Simon. *Cutting Through Water: Towards a Posthuman Theory of Educational Relationality*. Lund: Avdelningen för pedagogik. Sociologiska institutionen. 2015. (Disp 2016).

Gómez, Maria del Carmen. *Teachers' Assessment and grading Practices in upper secondary Science Classrooms in Sweden. The Teachers' and Students' Perspectives*. Lund: Avdelningen för pedagogik. Sociologiska institutionen. 2017. (Disp 2017).



*My unicorn and I forged a friendship that was
Based a bit on love, based a bit on truth
He would capture songs with his indigo horn
He knew how to share them with others,
That was his vocation (Silvio Rodríguez)*

Good teaching and fair grading in natural sciences for young people is the area of interest in this thesis. The author is herself a natural scientist, teacher and pedagogue.

Every human has a desire to learn, to satisfy her curiosity and to feel the joy of insights and deeper understanding of our world. In the natural sciences, this may be accomplished by studying nature and its phenomena, organization and laws. Then, by logical reasoning, experiments and assessing the outcome, continuously abandoning and formulating new hypotheses, one can make nature, the basis for human life, more understandable.

A remarkable example of this struggle for truth, this lifelong learning, is Charles Darwin. He devoted his life to the constant search for knowledge. In open, fair discussion, by pure logical reasoning and evidence, he challenged the beliefs of his time, convincing his opponents and opening their eyes to the wonderful world of ours. In honour of true science, fair discussions and good pedagogy, the study of Charles Darwin is depicted on the cover of this dissertation.