Teaching Traditions in Science Teachers' Practices and the Introduction of National Testing

Accepted for publication in Scandinavian Journal of Educational Research, 2017

Malena Lidar*, Department of Education, Uppsala University, Box 2136, SE-75002 Uppsala, Sweden, malena.lidar@edu.uu.se

Martin Karlberg, Department of Education, Uppsala University, Box 2136, SE-75002 Uppsala, Sweden, martin.karlberg@edu.uu.se

Jonas Almqvist, Department of Education, Uppsala University, Box 2136, SE-75002 Uppsala, Sweden, Jonas.almqvist@edu.uu.se

Leif Östman, Department of Education, Uppsala University, Box 2136, SE-75002 Uppsala, Sweden, leif.ostman@edu.uu.se

Eva Lundqvist, Department of Education, Uppsala University, Box 2136, SE-75002 Uppsala, Sweden, eva.lundqvist@edu.uu.se

*corresponding author

Abstract

Our main interest in this article is to explore whether Swedish teachers changed their teaching and assessment practices in relation to the new national tests in science education that were introduced 2009. Data was collected using a web-distributed questionnaire, which was answered by 407 teachers. The concept of teaching traditions is used to capture patterns of what is emphasised by teachers in terms of goals and content in teaching and the design of the questionnaire was based on the concept of curriculum emphases. The results show two distinct groups of focus, which is compared with two traditions within science education: the Academic and the Moral tradition. The main content where teaching have been changed is in making science more applied than before, where applied not only mean the application of science knowledge to practical technical issues but also to moral and political issues.

Keywords: Teaching traditions, science education, standardized tests, teacher survey, educational change

Introduction

Even though every teacher is responsible for shaping the teaching situation in each classroom, teaching is not idiosyncratic. Over time, there is a strong historical continuity in the way teachers teach (eg. Aikenhead, 2006; Solomon, 1998; Roberts 2007; 2011; Öhman, 2004). Teachers work in a context of habits and traditions that influence their selections of teaching content, but they also need to respond to new elements in their practices. New elements, such as a new testing regime, could possibly challenge and alter teaching practices. A closer look at the science education subjects show that there are a wide range of purposes and topics that teachers ought to address. The aim that *all* students would develop scientific knowledge in order to participate in community life, making every day, social, moral and political decisions has in recent years in many countries become an increasingly clear objective in science education. At the same time, there are goals in science education stating that students should be prepared for further studies, even though advanced studies in science are relevant only to a minority of students (Roberts, 2007; Levinson, 2010; Linder et al., 2011). The wide range of goals in science education may lead teachers to find it difficult to know what to focus on (cf. Ryder and Banner, 2011). Also, teachers' personal views of what constitutes 'good' science education are important since these views condition their experiences of policy reform (Ryder and Banner, 2013; Wallace and Priestley, 2011).

National testing in biology, physics and chemistry in Year 9 of Swedish lower secondary schools were introduced from 2009. Year 9 is the final year of compulsory school and the tests are mandatory for all Swedish students. The intentions for introducing these tests were to improve learning results in science education, to increase effectiveness by strengthening schools' monitoring of students' skills and to create a more equitable and fair assessment and grading for the students (Ministry of Education, 2008). Additionally, an explicit intention presented by the Ministry of Education was that the tests should function as models in assessment, teaching and learning by concretising the syllabi. However, international research shows that standardized tests do not necessarily affect all teachers' practices in the same ways (Au, 2007; Anderson, 2011) and that the enactment of reforms is often not entirely in line with what the policy developers intended (Kahle, 2007).

One way of interpreting the intentions behind the introduction of national tests in science education is that they are supposed to govern the ways teachers teach – or at least to interpret the curriculum – with the goal of more uniformity. However, how this goal is achieved is not obvious. Even though teachers nationally are working to meet the same goals in the curriculum, the emphasis in teaching can be different, leading to different approaches to teaching.

Aims and research questions

The overarching aim in this article is to explore if and how different groups of teachers perceive the need to change their teaching and assessment practices in relation to the introduction of new national tests in the science subject areas. This includes an investigation into how science teachers privilege different aspects of the content in teaching. Previous research has found different patterns in the selections of goals and content in science education (e.g. Fensham, 1988; Östman, 1995; Roberts 1982, 2007). We call the patterns of inclusion and exclusion of different teaching content teaching traditions. The idea is that if a specific pattern of inclusion and exclusion of educational goals and content is made by many teachers over a period of time, a teaching tradition has been formed. The patterns shown in previous research are results of studies of textbooks and curricula and not of views held by practicing teachers. The first aim is therefore to investigate if teachers' statements about their teaching correlate with different patterns of teaching. The following research questions are explored:

- Are there any patterns in science teachers' selections of goals and content?
- If there are patterns of teaching goals and content, what are the characteristics of those patterns?

Second, we aim to explore how teachers use the national tests in relation to their own teaching and assessment practices. If there are different patterns in teaching, it is not farfetched to assume that teachers will relate to the national tests in various ways. The second purpose is to investigate if there is a relationship between different teaching traditions and teachers' perceived needs to change their teaching in accordance with national tests. We will investigate:

- To what extent do teachers claim to change their teaching as a consequence of national tests in science?
- What content is addressed when changes in teaching are described?

Finally, we will discuss potential consequences of the findings in relation to teaching traditions and the use of national tests as way of national governing of the teaching content.

Background and framework

Traditions in science education

Research on curriculum theory, curriculum history and the history of school subjects have highlighted tensions between different visions for education, and the research have shown that over time and in different places, it is possible to find patterns in the purposes of education and the selections made in instruction. These patterns make visible various purposes as well as different values and approaches in teaching and have been described in different ways.

For example, Fensham (1988) describes the discussion about what to include in science education as a conflict between two poles. On one hand, he writes, some stress that science education should introduce students to the world of science and the work of scientists. But others stress that teaching should focus on giving students functional scientific knowledge, that is, the skills they can use in everyday life.

Another way to understand the traditions in a subject is to analyse *curriculum emphases* (Roberts, 1982). The concept of curriculum emphases was introduced and used to identify and describe the regularity within the epistemological dimension in teaching. Analysing science syllabi and science textbooks, Roberts (1982) (in North America) and Östman (1995) (in Sweden) found seven different curriculum emphases in science education; *Correct explanation*, where the way to argue about why you should learn science is that scientists know what is good knowledge, which eliminates the need for education to problematize the epistemological dimension; *Solid foundation*, in which students need to learn about logic and cumulative aspects of knowledge; *Scientific skill development*, where the objective is for students to learn practical skills in the scientific craft; *The structure of science*, that involves teaching students the intellectual skills of the scientific craftsmanship; *Self as explainer*, that includes historical explanations and students' own explanations of scientific phenomena and events; *Everyday coping*, where knowledge is justified by that you can apply them to practical issues of everyday life and *Science, technology and decisions* in which education focuses aspects that have to do with applying scientific knowledge on moral or political problems. In a sense, curriculum emphases answer the students' question, "Why am I learning this?" The curriculum emphases are thereby a way of categorizing the goals of education.

The curriculum emphases can in turn be related to the two main Visions (I & II) in Western societies regarding how science education should be constructed in order to make the students "scientifically literate" (Roberts, 2007). Vision I is described as science reproducing its own products of concepts, laws, theories and methods. Vision II emphasises that education must include facts on a subject, but it must also include knowledge and skills that make the students able to use scientific knowledge in practical, existential, moral and political contexts (e.g. Roberts, 2011; Zeidler, 2003). An important part of the skills each student should learn addresses the epistemological dimension, but the important difference is that in Vision I it is the natural science activities that are highlighted, while in Vision II, the reciprocity between the epistemological dimension and the norms and values in science is emphasised (e.g. Zeidler, 2003). In relation to assessment, Roberts (2007), Orpwood (2001, 2007) and Aikenhead, Orpwood and Fensham (2011) point out a trend: it appears that the knowledge that Vision I prioritises that becomes tested.

In an analysis of the subject content, subject focus and curriculum emphases in textbooks, policy documents and professional development materials in Sweden between 1960-1990 Östman (1995, 1998) distinguish two distinct patterns: the Academic tradition and the Moral tradition[1]. The Academic tradition is quite similar to Roberts' Vision I, with the focus of teaching science so that students will be introduced into the knowledge of the scientific disciplines with the exception. The most important difference is that in the Academic tradition the purpose is *also* to teach students to be able to use this knowledge in order to understand and/or solve everyday phenomena and problems. Eggens (2004) show that it is reasonable to distinguish between two variations (positivist and constructionist) of the Academic tradition.

In the Moral tradition the main purpose for teaching science is to help students understand how to make informed decisions regarding moral and political questions with the help of science. In terms of curriculum emphases it is possible to find all seven within a Moral tradition, but Science, technology and decisions is the overriding emphasis. Within the Academic tradition this last emphasis is not found at all. The division between Vision I and Vision II is based on a difference between disciplinary knowledge and application of scientific knowledge, while the division between the Academic and Moral traditions is between the inclusion/exclusion of moral, ethical and political issues. Noteworthy is that application of scientific knowledge is included in the Academic tradition, whereas applications on value issues is part of the Moral tradition. See table 1 for a summary.

[1] In the Swedish text, Östman (1995) calls the selective traditions 'Akademisk' and 'Romantisk'. We chose to translate ´Romantisk' to Moral, since Moral more accurately describes the content of the tradition.

[Insert table 1 here]

The existence of teaching traditions in areas connected to science education practice has been investigated in some previous studies. In the area of environmental education and/or education for sustainable development, Öhman and Östman (2002) showed that virtually all teachers can be sorted into one of three environmental education traditions: fact-based environmental education, normative environmental education and education for sustainable development. This categorization is also validated by Borg, Gericke, Höglund, and Bergman (2012) and Sund (2016). In chemistry education, Van Driel, Bulte and Verloop (2008) investigated teachers’ beliefs about the goals of the chemistry curriculum in the Netherlands using curriculum emphases. They found that the chemistry teachers tended to support all curriculum emphases, but that there was a significant and substantial difference in the support given to the various emphases. The differences were foremost dependent on what levels of higher education students were preparing for.

The Swedish national tests

The national tests in science education were introduced at the end of a pedagogical period based on the curriculum from 1994.The curriculum from 1994 had a model of management based on objectives and results. The students should achieve certain goals combined with a competence-based view of knowledge (Wahlström & Sundberg, 2015, p. 25). The curriculum had objectives for students to achieve, which would give teachers a degree of local control over selecting teaching content and teaching practices. Many teachers experienced this approach as nebulous, and when Sweden got a new government in 2006, the curriculum was criticized for its lack of focus on students' knowledge. This criticism, together with Swedish pupils achieving worse results than before, both in national surveys and international comparisons (e.g. TIMSS and PISA), constituted arguments for the introduction of national tests in science education.

Since 2009 national tests are given in each of the major science subjects, biology, chemistry and physics. Schools are randomly assigned one of the tests, so every student only performs tests in one of the subjects. The content of the national tests is based on national policy documents, foremost the syllabus for each subject. The syllabi in all the science subjects contain objectives and criteria that deal with theoretical knowledge and practical laboratory work as well as the application of scientific knowledge in both everyday and political/moral contexts.

In analyses of the national tests given between 2009 and 2012, Almqvist & Lundqvist (2013) found that in order to pass the test, students were required *at least* to demonstrate knowledge privileged in education built on Vision I. For higher grades it was required that students can correctly answer questions about applications in science and ethical and moral issues, i.e. to use knowledge within Vision II. This means that in the education-for-all model, the tests only measure a part of what students, in accordance with the national curriculum for these subjects, are expected to know in Year 9.

Effects of standardized assessment

Research show that there are different ways that standardized tests influence teaching and learning practices (e.g. Linn, 2000). Au (2011) highlights that high-stakes testing is connected to norms about how and what to teach and also to questions about who should have control over classroom practices. Cimbricz (2002) notes that there is a clear relationship between standardised assessments, teachers' thoughts and teachers' experiences but emphasises that the relationship is complex and needs to be investigated.

The assumption that compulsory national tests affect teacher instruction and student learning has been both confirmed (e.g. Barnes, Clark and Stephens, 2000) and rejected (Cimbricz, 2002) in international research. Anderson (2011) concludes in a review article on test-based accountability that testing worldwide continues to increase, even though research studies on the effects of high-stakes tests do not necessarily show improved student learning outcomes. This review shows that science instruction is often narrowed because of the accountability pressure, and that teachers' teaching practices become more fact-based.

Another common effect of standardized tests is that the content of teaching is adapted to the content of the tests (Au, 2007; 2011; Hamilton and Berends, 2006). Au (2007) found that the main effects of this type of tests are that educational content is limited to what is tested, the subjects are fragmented into test-related pieces, and teachers increase their uses of teacher-centred pedagogy. However, this picture of the national tests is not one-sided. Au (2007) also found the opposite in a significant minority of the cases, that the teaching content was expanded, that knowledge was integrated, and that more student-oriented teaching was used. Grant (2001) shows that centralized testing affects teachers' teaching, but depending on various factors, especially the teachers' views on the subject content and learning, teaching is affected in different ways.

Donnelly and Sadler (2009) interviewed twenty-two science teachers, asking if and how they modified their teaching on the basis of standards, including high-stakes testing. They found that roughly half of the teachers described that they modified their teaching according to the standards, but the teachers approached the standards in different ways. For example, some teachers expressed negative views because they felt the standards were too broad; some focused on teaching to the tests; some felt that they already taught in line with the standards; and others saw the standards as a part of the ongoing cycle of educational change, and standards were seen as useful tools for improving practice.

There is reason to believe that the introduction of national tests can change teachers' perceptions of what counts as 'good' science instruction, effective teaching and effective assessment practice. However, it is not certain that all teachers' instruction approaches are affected in the same way.

Method

In the present study, the educational patterns identified in curriculum theory research were measured quantitatively with a group of science teachers currently working in Swedish lower secondary schools. A questionnaire was used to explore teachers' teaching practices, including their teaching goals, their choice of teaching content and the teaching methods used in the classroom, and their the assessment practices. The questionnaire was developed using previous research about curriculum emphases to explicate different aspects of teaching goals and content selection. Besides questions of goals and content, the questionnaire also contained background questions, questions about the selection of methods for teaching, and teaching materials, student participation and assessment and questions about how the teachers register that they have been affected by the new national tests. The survey questions are presented in conjunction with the results.

The questionnaire contained 43 questions/positions, and the primary form of questioning was through closed-form response alternatives. A number of free-text response alternatives were also included. The questionnaire was designed in collaboration with the research group. Several measures were taken to ensure that the questions were accessible and recognisable before distributing the survey, e.g. trials by one group of faculty members and a second group of practicing teachers.

In the first step of the analysis we focus on how teachers, in the survey questions, privilege among different goals and contents in teaching in order to investigate if teachers teaching correlate with different teaching traditions. This is done by an exploratory factor analysis that shows if different subscales (i.e. patterns in teachers' selections) can be found. The next step is (if different subscales are found) to find out if there are different aspects characterising these subscales (i.e. if the patterns include more characteristics in addition to goals and content). Groups of teachers that in their answers aligned more with one subscale than the other were then identified. Using significance testing we investigate if these two groups have common characteristics in terms of gender and teaching experience and if they have common preferences in their selections of methods of work, teaching materials, student participation and assessment practice. A more detailed description of this approach in the analyses is presented in the Results section.

To investigate if there is a relationship between different teaching traditions and teachers' perceived needs to change their teaching in accordance with national tests, we do descriptive analyses from the following questions: 'Do you think that what you teach is different from what is tested on the national tests?' and 'To what extent do you think your teaching has changed as a consequence of the introduction of national tests?'. To characterize the changes, a thematic analysis of the free-text answers to these same questions was performed (Braun and Clarke, 2006). Finally, using significance testing, we investigated if teachers who privilege different aspects of teaching according to the results from step one responds to the introduction of national tests in science education in different ways.

Participants and procedure

A web-based distribution was used. E-mail addresses were mainly collected through the schools' websites but also through e-mail or telephone contact with heads of schools and/or school administrators. The survey was distributed to a total of 1003 science teachers. The total sample consisted of two groups where one group (*n*=480) consisted of science teachers in grades 7-9 in 16 municipalities in the area around our university. The second group (*n*=523) consisted of teachers from 31 (out of 290) randomly selected municipalities across the entire country. The sample included representatives from metropolitan municipalities, city-sized municipalities, and villages, and these municipalities were geographically distributed throughout the country. This design of the two different study groups was intentional: it gave the opportunity to follow-up the survey with individual interviews in our local area (cf. Lundqvist and Lidar, 2013; Sund, 2016), while allowing for measurements of representativeness.

The use of two sample groups was not necessary for the research presented here, and we wanted to be able to regard the two as one coherent group. A comparative analysis of 13 questions concerning the goals and content in teaching as well as 13 questions concerning other aspects of teaching was conducted. The *t*-tests showed that there was only one significant difference between the two groups. The teachers in the second sample group more often planned the lessons together with their students [$t(387) = -2.00$, $p < .05$]. One significant difference is to be expected with $p<.05$. Since the division in two groups has no significance for the research questions in the present study, hereafter, the groups will be considered as one coherent group.

The survey was sent by e-mail, which included a link to a web page. A first reminder was sent six days later and a second reminder a week after that. The survey was open for 30 days. The survey administrator received nine e-mails from persons who received the questionnaire but claimed they did not belong to the target group. In the local study group 15 were not contactable, 27 were unregistered, 251 did not respond, and 191 answered. In the national study group 16 were not contactable, 24 were unregistered, 268 did not respond, and 216 responded. Data was automatically generated from the web survey tool to an Excel spread sheet. The Excel data was imported into the Statistical Analyses Software (SAS® 9.3. [Computer software]. Cary, NC, USA: SAS Institute Inc.) for data analysis. In total, 407 teachers answered the questionnaire, which gives a response rate of 41%. With this response rate, we will not claim to provide results that are possible to generalize to the population, but this sample still includes a great deal of information. Archer (2008) claims that the effort and costs to increase the response rate in a web-based survey that seeks to work with theoretical development and improvement is not justified. The information in the responses contains sufficient information to show if diversities in teachers' manners of teaching exist or not.

Results

Measures and analysis of teaching traditions

How teachers select goals and content in teaching was measured with 13 items on the questionnaire. The teachers were asked to value how important different goals and content were in their teaching. Each item was followed by a 5-point scale ranging from "not important" to "very important".

The questions about goals addressed how the respondents rated the importance of students being:

(1) well-rounded in science education,
(2) prepared for further studies,
(3) prepared for future professions,
(4) prepared to encounter science in everyday life and
(5) prepared to serve as active and critical citizens

The questions concerning educational content addressed how the respondents rated the importance of:

(1) facts, concepts, models and relationships
(2) applications in everyday life

(3) applications to questions of a technical nature
(4) applications to social and political issues / problems
(5) applications to questions of existential and moral character
(6) working methods
(7) ways of thinking and reasoning
(8) the development and growth of knowledge

An exploratory factor analysis (EFA with Varimax rotation) was conducted in order to investigate the patterns among the collected answers (n=378). The data was suitable for factor analysis (Kaiser's MSA=0.81). The EFA revealed four factors, but whereof three factors without satisfactory internal consistency (they had a Cronbach's alpha below .70). In order to improve the internal consistency and interpretability, two-factor and three-factor solutions were examined.

The analyses showed that the two-factor solution gave the subscales with the best internal consistency. This means that the analysis has clarified two distinct patterns of selection of educational goals and contents made by teachers, in other words, what we will call two different teaching traditions.

The two-factor solution also offered a desirable interpretability, motivated on theoretical accounts, both by Roberts' (2007, 2011) use of two different visions for science education and Östman's (1995, 1996) results in terms of traditions. When compared to table 1 we see that Vision I and II are represented in factor 1 and 2 but also represent the Academic tradition and the Moral tradition respectively. Our conclusion is that the latter interpretation is preferable since the two items– "Applications to questions of a technical nature" and "Applications in everyday life" are found in factor 1 and applied science is a part of the Academic tradition, but is not present in Vision I.

In the two-factor solution the factor loadings were satisfactory (.30 or greater, see for example Raykov and Marcoulides, 2008). The first factor accounted for 75 % of the variance in the data. The internal consistency was satisfactory for both scales (*Cronbachs alpha .75; .72*) and the item-to-total correlations were satisfactory (.30 or greater, see for example Ferketich, 1991), see table 2.

[Insert table 2 here]

An additional explorative factor analysis was performed on each factor, on theoretical accounts. Analysis of the items within factor 1 gave two new subscales with an Eigenvalue over 1 (3.00; 1.09). Table 3 shows the results from a second-step exploratory factor analysis with a fixed number of two factors.

[Insert table 3 here]

Factors 1a and 1b could be compared to Eggen's (2004) division of the Academic tradition into a positivist and a constructivist branch. However, Nunnally and Bernstein (1994, p. 265) claim that the internal consistency should be above .70. In the second-step analysis, none of the factors reach this threshold. Consequently, we kept factor 1 intact for subsequent analyses.

Further analysis of the included items in factor 2 show that the four items should not be separated; there was only one factor with an Eigenvalue over 1 (2.19)(n=391)(Kaisers MSA 0.69). The two factors constitute two subscales, which were used for the continued analysis of the characteristics of the two subscales.

It can be concluded that there is evidence that teachers in their answers privilege differently among aspects of goals and contents and the subscales correspond well with the Academic and Moral tradition identified by Östman (1995).

In order to further explore the different subscales and identify if there were aspects that could be seen as more characteristic for one subscale than the other, χ^2-tests were performed. Since the data for subscale 2 were non-normal and had a kurtosis value above an absolute value of two, non-parametric tests were used in the comparison (Raykov and Marcoulides, 2008).

A χ^2-test examines if the frequencies of a number of different outcomes are similar to the hypothesis of a given probability distribution. The hypothesis in this case is that the probability distribution is the same in both subscales. To conduct the χ^2-tests we needed to select the respondents who in their answers had valued the items in one subscale higher than they valued the items in the other. The respondents' values of the items of each subscale were summed and the difference between the item sum of subscale 1 and subscale 2 were calculated. The respondents who answered with differences below the first quartile and above the third quartile were selected to uncover the groups whose answers aligned the most with subscale 1 and 2 respectively.

In comparisons the answers in each group of the background questions, it was clear that there were no significant differences between respondents in the different subscales concerning gender [χ^2 (df=1) = 1,47, p = n.s.]. In total 63% of the respondents were women and 37 % were men (n=404). Additionally, there were no significant differences in how teachers privileged among aspects of goals and content in relation to the amount of teaching experience they had [χ^2 (df=2) = .56, p = n.s.].

Consequently, we can conclude that aspects such as gender and teaching experience were not related to the two subscales. For example, women do not claim to be teaching more in alignment with specific goals and content, nor do experienced teachers assert to favour certain ways of selecting teaching content.

Teachers' approaches to teaching were examined through questions about methods of work in teaching, including lessons with briefings and tasks, investigative approaches, or discussions about the role of science in society. Teachers' use of teaching materials were measured by questions of the use of different information sources in instruction, namely textbooks, the Internet, field trips, printed business information, external lecturers, newspapers and popular science magazines. Finally, one question aimed at measuring teachers' positions on students' influence over the working processes, including the teacher plan without the influence of students, when students are responsible for planning under teacher supervision, and when the teacher collects opinions from students that are taken into consideration in the planning for teaching.

The teachers were asked to estimate the extent to which they used these different approaches in their teaching. The response alternatives were rated on a 5-point scale, in the first two questions based on lesson frequency (1=every lesson 2=once a week, 3=once a month, 4=once a semester and 5=never), and in the last question referring to more general frequency (1 = always, 5 = never). The predictor items were dichotomized in order to facilitate continued processing and analysis, based on a discussion about reasonable differences in practice and the demand of having sufficiently large samples. The probability distributions are presented in table 4.

[Insert table 4 here]

The teachers who scored high in subscale 1 were significantly less prone to use "lessons with briefings and tasks", compared to the teachers who scored high in subscale 2. The teachers who scored high in subscale 2 more often engaged the students in "discussions about the role of science in society" compared to the teachers who scored high in subscale 1, whereas there were no significant differences in how often these teachers used investigative work in their teaching. In the question about the use of teaching materials, the significant differences were that teachers who scored high in subscale 2 reported that they used the Internet, business information and popular science magazines more frequently than teachers who scored high in subscale 1. In the question about student participation, the teachers who scored high in subscale 2 to a higher degree encouraged their students to work thematically and in groups and let their students be responsible for planning under their supervision. It is interesting to note that the respondents whose answers aligned with Subscale 2 represented the group with more frequent use of all the above items.

Responses to national tests, in relation to teaching traditions

The second aim was to investigate if there is a relationship between different teaching traditions and teachers' perceived needs to change their teaching approaches in accordance with national tests. The first step to explore this aim is to investigate to what extent teachers state that their teaching is different from what is tested on the national tests and to what extent they describe that they have changed their teaching. This was measured with the questions 'Do you think that what you teach is different from what is tested on the national tests?' and 'To what extent do you think your teaching has changed as a consequence of the introduction of national tests?'.

The results show that as many as 90% of all teachers answered that they have changed their teaching to some extent. Almost half (44%) of all the responding teachers who have administrated the national tests (n=339) rated their teaching as different from what is measured on the national tests. The most frequent answer was that their teaching had changed to a small degree, but approximately one third (31%) rated that their teaching had changed quite a lot or to a high degree since the introduction of national tests (n=339).

A significant change is undoubtedly present. In order to understand how different teachers described the change, we take a closer look at the free text comments in conjunction to the questions about differences between the national test and teachers' teaching. In the comments field, the teachers were asked to describe *if* they thought their teaching is different, *how* it is different, and *if* their teaching has changed, *how* it has changed.

Altogether, 230 comments were written for the two questions. All statements were coded in an iterative process of reading and re-reading and then categorized into themes that include descriptions of changes in content selection in teaching. Out of the 230 comments, 138 deals with differences between the tests and the teachers' own instructional decisions in terms of teaching content and was included in the analysis. Among the excluded answers are answers that do not relate to teaching content and answers that are difficult to interpret. For example, when words like "whole" and "connections" are used without elaboration, it is difficult to understand what the changes in teaching content the comment refers to.

The remaining answers concern two main themes. The first theme focus on the view that the respondent's instruction is more focused on facts and laws than the tests are (theme 1a, 11 answers) or that the tests are more focused on societal issues (environmental issues, sustainable development, everyday questions etc.) than their instruction (theme 1b, 49 answers). An example of answers in within theme 1a reads,

> [...] it feels as if I'm asking for a little more "factual knowledge" than the national test is looking for. Students who are good with good basic knowledge expressed that they thought the national test contained mostly a lot of "general" questions and not so much specific knowledge of chemistry. On the other hand, I could deduce that these students could more clearly support the foundation of their "opinions" and answers, and in that way they also showed their subject knowledge. (Q78)

An example of theme 1b is found in answers where a teacher states that, in contrast to his or her own teaching, the test requires students to identify which arguments or reasoning that are based on scientific arguments (rather than economic or emotional arguments). One respondent wrote,

> Yes, I have become more influenced by the national test. I hadn't previously thought (understood?) that so much emphasis should be put on areas such as, "What is based on scientific opinions?" (Q115)

Another expression of this category is when teachers describe their teaching as including more traditional disciplinary (academic) science than the national tests emphasise, with more fact-based teaching and, in some cases, formula writing and calculations. For example, teachers stated,

> The national test in chemistry had more focus on cycles and sustainable development, etc., and less focus on chemistry models and theories, compared to my teaching. (Q69)

> My teaching reveals more of the nature of the subjects while the national tests lean towards more social- and nature-oriented knowledge (national tests often do not feel subject-specific). (Q75)

Our interpretation of this is that teachers acknowledge that there is a discrepancy between the completed instruction and the tests. In other words, the traditional academic content is more frequently used in their instruction than in the tests. If we couple this with the fact that 90% of the teachers perceive that they have changed their teaching to some extent, it would seem likely to conclude that the Academic tradition will decline in importance and that future instruction will contain more societal elements, such as making arguments for the use of different energy sources. In other words, the Moral tradition will have a greater impact.

The results could also be interpreted somewhat differently. What distinguishes the Moral tradition from the Academic is that questions of moral and political character are addressed in the former; for example, the students should learn to make arguments with scientific knowledge and take a stand on societal questions where there is disagreement. In both teaching traditions, instruction leading to learning about the concepts, laws and theories of science is important, as is teaching about the applications of this knowledge in order to solve everyday or technical problems. The free text answers in the questionnaire could therefore be interpreted as if teachers perceive that their instruction is more directed towards teaching the students scientific knowledge, while the tests contained more applied science. However, whether teachers claim that these applications are related to questions of everyday or technical character, of moral or political character or of both is not possible to say for sure. Nevertheless it is clear that the teachers’ responses can be understood as expressions that the tests, in contrast to instruction, for some represent a shift in favour of the applied teaching in the Academic tradition and/or to the benefit of the Moral tradition.

If we address comments only from the respondents who have been identified as aligned mostly with either of the traditions, it is apparent that comments stating more use of disciplinary science is more frequent in the group aligned with Academic tradition. In the group mostly aligned with the Academic tradition, (n=45) we received 31 responses, whereof 10 are claims for this stance. Among the teachers mostly aligned with the Moral tradition (n=38), we received 19 responses, whereof only one mentions that he/she includes more teaching of facts, but in this case at the expense of investigative work (Q206).

Our interpretation is thus that teachers, who in their answers align more with the Academic tradition experience discrepancies in the content between their own teaching and the tests concerning elements of learning when applying scientific knowledge to solve political and moral problems in society, which teachers answering in alignment with the Moral tradition do not experience to the same extent.

This conclusion is understandable in comparison to analyses of the test items: Almqvist & Lundqvist (2013) show that some of the test data can be categorized within what we here call the Moral tradition. However, this reasoning does not exclude the possibility that teachers' responses also involve a shift in favour of applied teaching in the Academic tradition. If we are to believe that teachers will change their teaching when working with national tests, we could conclude that the Academic tradition diminishes in importance in favour of the Moral tradition but also that the applied component - the application of scientific knowledge to everyday technical problems - in the Academic tradition also will have a significant role.

The second theme concerns the idea that the tests present a specific way of doing laboratory work; 55 responses report this. Teachers write that they are now working more with letting their students plan their investigations themselves, evaluate their plans and results and find sources of error from the performance of the lab. These answers were not correlated to alignment with the Academic and the Moral tradition though. Yet another theme that is mentioned is that the tests include more of the history of science than in instruction (8 claims). Our conclusion is thus that the analyses show that there seems to be changes in the teaching content that possibly could direct teachers towards teaching more in alignment with a Moral tradition.

Discussion

In this article we have explored if science teachers' statements about their teaching correlate with different teaching traditions and how they describe changes in their teaching and assessment practices in relation to the new national tests. We have identified patterns in how teachers privilege different aspects of teaching, and we have found similarities between how this is done and the patterns presented in previous research and theoretical models. The study includes the limitation that what we measure is teachers' statements about their practice, and not their actual practice. Nevertheless, these statements tell us something about what they privilege and we can see that the reported historical patterns have parallels with how teachers privilege aspects of instruction today.

The patterns identified in our study correspond well to the Academic and Moral traditions found in Sweden as presented by Östman (1995, 1996). The items included in Subscale I in the factor analysis are in alignment with the Academic tradition, where the respondents answered that they emphasise the teaching of concepts, theories and methods of science that could be seen as preparation for further studies in science, and they emphasise using scientific knowledge in everyday applications to a larger extent. The items included in subscale II correspond with the description of the Moral tradition since the focus here is more on existential, political and moral applications of scientific knowledge. We found that these differences were not related to gender or experience of teaching. However, there were some differences in the groups of teachers that answered mostly in alignment with the different traditions. Lessons with briefings and tasks, discussions about the roles of science in society, the use of the Internet, business information and popular science magazines, thematic work and students' involvement in the planning of teaching are aspects that are more frequently highly rated by teachers answering in alignment with the Moral tradition.

We also examined if teachers who answered in alignment with the different teaching traditions responded to the introduction of national tests in different ways. However, this difference could not be confirmed on a statistical level. When selecting the teachers aligning the most with either the Academic or the Moral tradition and comparing if they considered their teaching to be in alignment with the national tests, there were no significant differences between the groups. However, this result only addresses the degree of change, and not the content of the perceived change. From analyses of the free text comments we get information that in the perceived differences between teachers teaching content and the content in the tests there is a tendency towards that teachers acknowledge a difference where more content of applied and moral character is included in national tests than in their own teaching.

With a point of departure in earlier research regarding the function of national tests showing that high-stakes tests often do function as prescribed curriculum (e.g. Au 2007, 2011) there is a possibility that science teaching in Sweden will be more applied than before, applied not only in the meaning of application of science knowledge to practical technical issues but also to moral and political issues. In some of the written comments given by the teachers regarding the similarities and differences between their own teaching content and the content of the test, we can find a suggestion that the national tests make teachers that privilege the Academic tradition more inclined to incorporate additional content reflecting the Moral tradition in their practices. The results concerning changes in relation to the Moral tradition are more difficult to interpret, but one area of content that seem to overarch all traditions is how the lab work is performed.

Consequently, from our analyses we conclude that the way the national tests are approached could possibly lead to a displacement toward a more applied science education. However, in an Academic tradition "applied" refers to practical and technical questions and problems in everyday life, while in the Moral traditions “applied” refers also to existential, moral and political questions and problems in everyday life. The difference is huge and it is difficult to determine which direction a potential change will take. More research is necessary in order to get the full picture.

Maybe the most important finding in this study is the fact that we have found patterns regarding teachers’ perceptions of the content they teach in science subjects, and that these patterns can be compared to the teaching traditions earlier historical research has found. Thus, Academic and Moral traditions seem to be valid. To have the knowledge of existing teaching traditions creates the possibility to extract generalisations out of qualitative data focusing on individual teachers. The knowledge gained from studies on individual teachers runs the risk of becoming anecdotal. With the knowledge of teaching traditions we can make more generalised claims and generate more generalised knowledge. This knowledge can be used, for example, when designing studies in order to guarantee coverage of a variety of manners of teaching. The knowledge of teaching traditions may build an important foundation in smaller qualitative studies, as a general survey can provide the basis for selection.

If it is the case that teachers aligning with different traditions of teaching understand a reform in different ways, this will have consequences for how policy makers need to work to effectively implement reforms. For example, it might be helpful to use different study materials or professional development for different groups of teachers in order to influence them in the desired direction, to appeal to teachers who, on the one hand, align themselves with the Academic tradition and, on the other, with the Moral tradition. Through knowledge of selective traditions and how they seem to work, we have the possibility to better understand teachers' reactions to new elements in teaching. The traditions can be used as tools with which we can understand the new reforms, without having to rely solely on qualitative studies. However, further studies are needed to determine the extent to which the different teaching traditions occur.

Acknowledgement

This work was supported by the Swedish Research Council under Grant UVK 2010-5153 and UVK 2012-5023. The final work on the article was done during an intensive working week on Fårö, Gotland - thanks to the Ingmar Bergman Estate Foundation, which made this possible.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Aikenhead, G. S., Orpwood, G., & Fensham, P. (2011). Scientific Literacy for a Knowledge Society. In C. Linder, L. Östman, D. A. Roberts, P.-O. Wickman, G. Erickson & A. MacKinnon (eds). *Exploring the Landscape of Scientific Literacy*. pp. 28-44. New York and London: Routledge.

Almqvist, J. & Lundqvist, E. (2013). De nationella provens innehåll: Vilken scientific literacy mäts i NO-proven. [The content in the national tests: What scientific literacy is measured in the science tests?] In E. Lundqvist, R. Säljö and L. Östman (eds) *Scientific literacy,* pp. 101-117, Malmö: Gleerups Utbildning AB.

Anderson, K. J. B. (2011). Science Education and Test-Based Accountability: Reviewing Their Relationship and Exploring Implications for Future Policy. *Science Education, 96:* 104-129.
Archer, T. M. (2008). Response Rates to Expect from Web-Based Surveys and What to Do About It. *Journal of Extension*, *46*(3). Retrieved from http://www.joe.org/joe/2008june/rb3.php.
Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, *36:* 258-267.
Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies, 43*(1): 25-45.
Barnes, M., Clarke, D., & Stephens, M. (2000). Assessment: The engine of systematic curricular reform? *Journal of Curriculum Studies, 32:* 623-650.
Borg, C., Gericke, N., Höglund, H.-O. & Bergman, E. (2012). The barriers encountered by teachers implementing education for sustainable development: Discipline bound differences and teaching traditions. *Research in Science & Technological Education, 30*(2): 185-207.
Braun, V., & Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2): 77-101.
Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practice. *Education Policy Analysis Archives*, *10*(2). Retrieved [150101] from http://epaa.asu.edu/epaa/v10n2.html.
Donnelly, L. A., & Sadler, T. D. (2009). High School Science Teachers' Views of Standards and Accountability. *Science Education, 93:* 1050-1075.
Eggen, A. B. (2004). *Alfa and Omega in student assessment. Exploring identities of secondary school science teachers*. Oslo: Department of Teacher Education and School Development, University of Oslo, Norway.
Englund, T. (1986). *Curriculum as a political problem. Changing educational conceptions, with special reference to citizenship education*. Lund: Studentlitteratur.
Fensham, P. J. (1988). Familiar but different: some dilemmas and new directions in science education. In P. J. Fensham (ed.). *Development and dilemmas in science education*, pp. 1-26. London: Falmer Press.
Fensham, P. J. (2009). The link between policy and practice in science education: The role of research. *Science Education, 93:* 1076-1095.
Fensham, P. J. & Corrigan, D. J. (1994). The implementation of an STS chemistry course in Australia: A research perspective. I J. Solomon & G. Aikenhead (eds). *STS Education: international perspectives on reform*. pp. 194-204. New York: Teachers College Press.
Ferketich, S. (1991). Focus on psychometrics: Aspects of item analysis. *Research in Nursing & Health*, *14:* 165–168.
Grant, S. G. (2001). An uncertain lever: Exploring the influence of state-level testing on teaching social studies. *Teacher College Record, 103:* 398-426.
Hamilton, L. S., & Berends, M. (2006). Instructional practices related to standards and assessments. *RAND*. Retrieved from http://www.rand.org/content/dam/rand/pubs/working_papers/2006/RAND_WR374.pdf
Kahle, J. B (2007). Systematic reform: Research, vision, and politics. In S. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 911-942). Mahwah, NJ: Erlbaum.
Linder, C., Östman, L., Roberts, D. A., Wickman, P. O., Erickson, G., & MacKinnon, A. (Eds.)(2011). *Exploring the landscape of scientific literacy*. London: Routledge
Linn, R. L. (2000). Assessment and accountability. *Educational Researcher*, *29:* 4-16.

Levinson, R. (2010). Science education and democratic participation: an uneasy congruence. *Studies in Science Education, 46*(1): 69-119.
Lundgren, U. P. (1979). *Att organisera omvärlden - en introduktion till läroplansteori*.[To Organize the World - An Introduction to Curriculum Theory.] Stockholm: Liber.
Lundqvist, E., Almqvist, J., & Östman, L. (2012). Institutional traditions in teachers' manners of teaching. *Cultural Studies of Science Education, 7:* 111-127.
Lundqvist, E., & Lidar, M. (2013). Nationella prov i NO och lärares val av undervisningsinnehåll.[National Tests in Science Education and Teachers' Selection of Teaching Content] *Utbildning & Demokrati, 22*(3): 85-106.
Ministry of Education (2008). Departementspromemoria. Fler obligatoriska nationella prov i grundskolan m.m. [Departmental memorandum. More Compulsory National Tests in Compulsory Education, etc.] Retrived [2008-10-10] from http://www.regeringen.se/content/1/c6/10/82/84/8c8adb7a.pdf
Munby, H., & Roberts, D. A. (1998). Intellectual independence: A potential link between science teaching and responsible citizenship. In D. A. Roberts & L. Östman (eds), *Problems of meaning in science curriculum*, pp 101-114 New York: Teachers College Press.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory*, 3rd ed. New York: McGraw-Hill.
Öhman, J. (2004). Moral perspective in selective traditions of environmental education. In *Learning to change our world*, In P. Wickenberg, H. Axelsson, L. Fritzén, G. Helldén, and J. Öhman (eds), pp 33–57. Lund: Studentlitteratur
Öhman, J., & Östman, L. (2002). *Miljöundervisning och utbildning för hållbar utveckling i svensk skola.* [Environmental Education and Education for Sustainable Development in Swedish Schools] Stockholm: Skolverket.
Orpwood, G. (2001). The role of assessment in science curriculum reform. *Assessment in Education, 8*(2): 135-151.
Orpwood, G. (2007). Assessing scientific literacy: Threats and opportunities. In C. Linder, L. Östman and P.-O. Wickman (eds). *Promoting scientific literacy: Science education research in transaction. Proceedings of the Linnaeus Tercentenary Symposium* (pp. 120-129). Uppsala: Uppsala University.
Östman, L. (1995). Socialisation och mening: No-utbildning som politiskt och miljömoraliska problem. *Uppsala Studies in Education, 61*. Stockholm: Almqvist & Wiksell.
Östman, L. (1996). Discourse, discursive meanings and socialization in chemistry education. *Journal of Curriculum Studies, 28:* 37-55.
Östman, L. (1998). How companion meanings are expressed by science education discourse. In D. A. Roberts and L. Östman (eds). *Problems of meaning in science curriculum* (pp. 54-70) New York: Teachers College Press.
Raykov, T., & Marcoulides, G.A. (2008). *Introduction to applied multivariate analysis.* New York: Routledge, Taylor and Francis Group.
Roberts, D. A. (1982). Developing the concept of "curriculum emphases" in science education. *Science Education, 62,* 243-260.
Roberts, D. A. (1988). What counts as science education? In P. J. Fensham (ed.), *Development and dilemmas in science education,* pp. 27-54. New York: Falmer Press.
Roberts, D. A. (2007). Scientific literacy/science literacy. In S. K. Abell & N. G. Lederman (Eds.). *Handbook of research on science education,* pp. 729-780. Mahwah, NJ.: Lawrence Erlbaum.
Ryder, J., & Banner, I. (2011). Multiple aims in the development of a major reform of the national curriculum for science in England. *International Journal of Science Education, 33:* 709-725.

Ryder, J., & Banner I., (2013). School teachers' experiences of science curriculum reform. *International Journal of Science Education 35:* 490-514.
Solomon, J. (1998). The science curricula of Europe and the notion of scientific culture. In D. A. Roberts & L. Östman (Eds.), *Problems of meaning in science curriculum*. New York: Teachers College Press.
Sund, P. (2016). Discerning selective traditions in science education:a qualitative study of teachers' responses to what is important in science teaching. *Cultural Studies of Science Education.* 11: 387. Van Driel, J. H., Bulte, A. M., & Verloop, N. (2008). Using the curriculum emphasis concept to investigate teachers' curricular beliefs in the context of educational reform. *Journal of Curriculum Studies, 40*(1): 107-122.
Wahlström, N. & Sundberg, D. (2015). Theory-based evaluation of the curriculum Lgr 11. Working paper 2015:11, Uppsala: Institute for Evaluation of Labour Market and Education Policy IFAU.
Wallace, C., & M. Priestley (2011). Teacher beliefs and the mediation of curriculum innovation in Scotland: A socio-cultural perspective on professional development and change. *Journal of Curriculum Studies*, 43, 357–381.
Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and Web survey services. *Journal of Computer-Mediated Communication*, *10*(3), article 11.
Zeidler, D. (ed.) (2003). *The role of moral reasoning on socioscientific issues and discourse in science education*. London: Kluwer Academic Publishers.

Table 1. The relations between teaching content, in terms of curriculum emphases, and educational traditions. Bold is used to mark the main goal/s within a tradition. The other curriculum emphases are subordinate and may or may not be a part in teaching in order to reach the goal.

Selective traditions (Östman 1995)	Academic - positivistic	Academic - constructivist	Academic - applied	Moral
Visions of science education (Roberts 2007, 2011)	Vision I	Vision I	Vision II	Vision II
Curriculum emphases (Roberts 1982; Östman 1995)	**Correct explanation**	Correct explanation	Correct explanation	Correct explanation
	Solid foundation	Solid foundation	Solid foundation	Solid foundation
	Scientific skill development	**Scientific skill development**	Scientific skill development	Scientific skill development
	Structure of science	**Structure of science**	Structure of science	Structure of science
		Self as explainer	Self as explainer	Self as explainer
			Everyday coping	Everyday coping
				Science and decisions

Table 2: The two-factor solution for the scale with rotated factor loading, aspects of goals and contents

Factor 1			Factor 2		
Cronbachs alpha: .75 (n=391-398)	Factor loading	Item to total correlation	Cronbachs alpha: .72 (n=392-398)	Factor loading	Item to total correlation
Working methods	.66	.44	Applications to social and political issues / problems	.78	.54
The development and growth of knowledge	.62	.53	Applications to questions of existential and moral character	.76	.56
Prepared for further studies	.60	.36	Prepared to serve as active and critical citizens	.69	.56
Applications to questions of a technical nature	.53	.44	Prepared to meet everyday life	.62	.39
Facts, concepts, models and relationships	.53	.35			
Ways of thinking and reasoning	.51	.42			
Prepared for future professions	.48	.43			
Applications in everyday life	.45	.48			
Well-rounded in science education	.35	.32			

Table 3. Second-step two-factor solution for the scale with rotated factor loading, aspects of goals and contents

Factor 1a			Factor 1b		
Cronbachs alpha: .56	Factor loading	Item to total correlation	Cronbachs alpha: .69	Factor loading	Item to total correlation
Well-rounded in science education	.73	.34	Working methods	.70	.47
Facts, concepts, models and relationships	.73	.36	The development and growth of knowledge	.72	.54
Prepared for further studies	.53	.30	Ways of thinking and reasoning	.70	.42
Applications in everyday life	.48	.36	Prepared for future professions	.57	.40
			Applications to questions of a technical nature	.55	.42

Table 4. Probability distribution of working methods, teaching materials and student participation

Item	Dichotomization	Differences
Lessons with briefings and tasks	1/2	χ^2 (df=1) = 5.13, p < 0,05
Investigative approaches with e.g. laboratory work and excursions	2/3	χ^2 (df=1) = 3.43, p =n.s.
Discussions about the role of science in society	2/3	χ^2 (df=1) = 5.30, p < 0,05
Textbooks	1/2	χ^2 (df=1) = 1.87, p = n.s.
Internet	2/3	χ^2 (df=1) = 3.96, p < 0,05
Field trips	4/5	χ^2 (df=1) = 2.34, p = n.s.
Printed business information	4/5	χ^2 (df=1) = 6.77, p < 0,05
External lecturers	4/5	χ^2 (df=1) = .35, p = n.s.
Newspapers	3/4	χ^2 (df=1) = 3.72, p = n.s.
Popular science magazines	3/4	χ^2 (df=1) = 6.08, p < 0,05
I plan my teaching myself; students do not participate to any great extent	2/3	χ^2 (df=1) = 1.23, p = n.s.
Students work thematically and in groups and are responsible for planning under my supervision	3/4	χ^2 (df=1) = 6.22, p < 0,05
Students express views that I collect and use in my planning	2/3	χ^2 (df=1) = 1.16, p = n.s.

Biographical notes

Malena Lidar is an associate senior lecturer of curriculum studies at the Department of Education, Uppsala University, Sweden. Her research explores issues of teaching and learning in science education, with a focus on the selection of teaching content in relation to standardised assessment as well as with issues of power and knowledge in science and technology classrooms.

Martin Karlberg is a senior lecturer of curriculum studies at the Department of Education, Uppsala University, Sweden. Karlberg's research focuses primarily on classroom management and he is currently involved in a project concerning how pedagogical climate affects school achievement and mental health.

Jonas Almqvist is a professor of curriculum studies at the Department of Education, Uppsala University in Sweden. Most of his research has been concentrated on teaching and learning in science education classrooms, but also includes an interest in meaning making processes in physical education and health.

Leif Östman is a professor of curriculum studies at the Department of Education, Uppsala University. He is the leader of the research group SMED (Studies of Meaning Making in Educational Discourses) where pragmatism, post-structuralism and the work of Wittgenstein are the main sources for inspiration. His own research focuses on the interplay between intrapersonal, inter-personal and institutional dimensions of learning as well as the role of values in the educational practices and discourses.

Eva Lundqvist is a senior lecturer of curriculum studies at the Department of Education, Uppsala University, Sweden. Lundqvist's research interest is in science education. Her interest foremost concerns teachers' role for students' learning, but also how policy changes influences teachers' teaching and assessment practice.