

Adaptation and resistance: washback effects of the national test on upper secondary Swedish teaching

Maria Larsson^{a,*} and Christina Olin-Scheller^b

^a*University of Dalarna, Sweden;* ^b*Karlstad University, Sweden*

This study examines the washback effects of a national test in Swedish at upper secondary schools. The test consists of three parts—reading, writing and speaking—and this study specifically focuses on the washback effects in relation to Stephen Ball’s theory of policy enactment. The study draws on qualitative data from a total of 21 interviews conducted with five teachers of Swedish in upper secondary education in a series of four individual interviews and a concluding group interview. The results indicate that the reading comprehension test has had little washback effect on reading instruction whereas the writing and the oral tests seem to have had a substantial washback effect. The teachers also exhibit what Ball refers to as resistance and capitulation in relation to the national test, and to the reading comprehension test in particular.

Keywords: washback effects; policy; Swedish as a school subject; national tests; upper secondary school

Introduction

In their everyday practices, teachers constantly have to take into consideration various guidelines and recommendations—national and local as well as formal and informal documents. These guidelines and recommendations can also be described as different policies (Ball, 1993), and national tests are examples of national policies. When external standardised tests are introduced, the expectation is that teachers will change the way they teach and start doing things differently in the classroom (Alderson & Wall, 1993; Shohamy *et al.*, 1996; Lundahl, 2009). The effects of national tests on teaching, or the so-called washback effects (Alderson & Wall, 1993) on teaching practices, are in focus in this article. According to Alderson and Wall (1993, p. 117), washback occurs when ‘teachers and learners do things they would not necessarily otherwise do because of the test’.

In the past few decades there has been an international trend of increased testing in schools. Decentralised school systems tend to use external standardised tests in order to control teachers and what and how they teach (Shohamy *et al.*, 1996). However,

*Corresponding author. Department of Humanities and Media, University of Dalarna, Högskolan Dalarna, 791 88 Falun, Sweden. Email: malr@du.se

research shows that attempts to change teachers' classroom practices through the introduction of an external test have succeeded just as often as they have failed (Chapman & Snyder, 2000). It is therefore important to understand more about how these external standardised tests are received and perceived by those who are supposed to change their practice so that policymakers and researchers can know more about the effects of these tests. In this article we want to contribute to an international discussion about washback effects of external testing in schools by giving an example from the Swedish context of teachers' descriptions of how their practice has—or has not—been affected by national standardised testing.

Ball (1993) states that policy is constantly being interpreted and enacted at different levels of the education system, and that different school processes are governed by policy. While a policy is a practical intervention in the classroom, it does not tell teachers what to do but rather creates circumstances that limit the number of available options (Ball, 1993). While teachers are policy subjects, they are also active policy actors who interpret, translate and transform policy into functional classroom practices (Ball *et al.*, 2012). Teachers therefore have double roles. In this interpretation process, teachers can choose to let their teaching be affected in different ways by the national tests. Washback is a complex phenomenon influenced by many different factors, such as teachers' attitudes to teaching and tests, their academic backgrounds and further training needs, teaching styles and teaching experience, as well as communication between those who design the tests and those who have to implement them (Wall, 2000; Luxia, 2004; Watanabe, 2004). National tests and washback effects must therefore be placed in a larger context, and it is important to examine how external standardised tests interact with different factors in the teaching situation and how they affect instruction in the classroom.

A policy-based approach can be taken to different general washback effects of tests, not least in a situation when new national tests are introduced and teachers are expected to adapt their teaching practices. For the past 10 years Sweden follows an international educational trend with more focus on students' results and an increasing number of external standardised tests, referred to in Sweden as national subject tests. Here, the national subject tests fulfil more purposes than in most comparable countries. In addition to supporting equitable grading, they should also enhance the attainment of learning outcomes and function as quality assurance in schools at local and national levels. They should further concretise learning outcomes and grading criteria, bring about change in teaching practice by supporting the implementation of syllabi and develop teaching, while also serving a diagnostic purpose (Lundahl, 2009; Skolverket, 2014). For example, one of the aims of the national test in the school subject Swedish in upper secondary school is to make teachers concentrate more on reading comprehension (Palmér, 2011). Thus, one way of investigating how teachers implement policy in their everyday classroom practices is to study the washback effects of the test on Swedish teachers' reflections on the choices they make when teaching.

The national tests in Sweden are administered and marked by the individual teacher, who also is the person responsible for deciding what role the results of the tests should be given when awarding an individual student a final grade on a course. Thus, the individual teacher is autonomous in the responsibility of assessing a final grade based entirely on the teacher's own data, including the results from national

tests if the course comprises such a test, and in upper secondary school the final grade is then included in the student's school leaving certificate. Because of its many purposes and because the results of the test are included in students' final grades, the expectations of the impact of the test in Sweden are high and studying the washback effects of such a test could shed light on the phenomenon not only in Sweden but also in other contexts and countries too.

In Sweden, upper secondary education is not compulsory but almost all students (98%) study at an upper secondary school. Students attend school for 3 years and choose between different programmes that either prepare for further academic studies, such as the natural or social science programmes, or prepare for vocational qualifications in for example engineering or care work. Regardless of programme chosen, there are some basic subject courses that all students study, such as Swedish (L1), Mathematics, English (L2) and social and natural sciences. Each subject course is framed by a syllabus that states the aim of the subject, core contents to be taught and criteria for assessment and grading. The grades range from F-A, where F indicates that the student has failed the course and A is the highest grade. For the passing grades E-A there are written definitions of what a student must achieve in order to receive a specific grade, so-called knowledge requirements.

In this investigation we focus on the course in the subject of Swedish that all students study during their first year of upper secondary school. This course comprises a national test consisting of three parts (reading, writing and oral skills) the students take during three different days at the end of the school year. The class teacher is mainly responsible for marking and assessing the tests and the results of the tests are included in the final grade of the course for the individual student. Since the national test does not test all aspects or contents of the syllabus, it does not count as a final course test. However, as the test impacts on the students' final grade, it is considered an important measure of students' achievements.

Although there is a substantial body of international research on how external standardised tests affect teaching, relatively little research has been done from a Swedish perspective. As mentioned above, the number of national tests in the subject Swedish has increased radically since 2011 (Skolverket, 2017). Studies show that this has changed teachers' roles (Lundström, 2017), but as yet little research has been done on the implications of such changes for teaching practises. Even if Larsson (2018) in an earlier study based on teachers' perspectives, has touched on the relation between policy and national tests in upper secondary schools, Swedish studies tend to focus on younger students (cf. Persson, 2017; Wetterstrand *et al.*, 2017). This article, however, takes a policy-based approach to investigating the washback of national tests on the subject of Swedish in upper secondary school, centring on issues such as how teachers' understanding of the test can be described, and on how teachers relate to the contents of the test when teaching.

Washback as phenomenon

In this article we use Alderson and Walls (1993) definition of washback effects as cited above. Previous studies on washback have provided a rather fragmented picture of the

actual effects of external standardised tests on teaching (Bailey, 1996). External tests may be an efficient way to influence teachers' classroom practices even though there is no conclusive evidence that these tests do affect teaching as intended (Chapman & Snyder, 2000; Andrews, 2004). The washback effect may also differ according to context since teaching is influenced by many different factors, and sometimes teaching does not seem to be affected at all (Spratt, 2005). Several studies also show that politicians and policymakers seem to regard the tests as efficient tools to control the educational system and thus to control and change teachers' behaviour (Shohamy, 1993; Shohamy *et al.*, 1996; Andrews, 2004; Ferman, 2004; Polesel *et al.*, 2014). However, if teachers really do change the way they teach seems to a large extent depend on whether policymakers succeed in understanding teachers' conditions and opportunities to bring about a change (Chapman & Snyder, 2000). There is also a risk that teachers are not fully aware of to what extent the tests influence their teaching and that they start teaching to the tests because they feel compelled to improve their students' results (Hardy, 2015).

If the tests reflect syllabi well and teachers teach in accordance with the syllabus and there is close correspondence between the test and syllabus, there is positive washback (Popham, 2001; Andrews, 2004; Turner, 2006). The opposite, negative washback, occurs when the teacher, for example, emphasises the test to such an extent that teaching pays too little attention to aspects that are not tested, or when the test and syllabus do not correspond well and teachers start teaching to the test instead of teaching the content indicated in the syllabus. If teachers experience the test as contradicting their own teaching principles, but still feel compelled to teach to the test, tension may arise between pedagogical and ethical decisions (Bailey, 1996).

In a Swedish context, Lundqvist and Lidar (2013) have investigated secondary school teachers' descriptions of how the national tests in science subjects affect the chosen teaching contents. Lundqvist and Lidar describe both positive and negative washback. Negative effects include increased workload, time pressures and a feeling of being governed. The positive effects highlighted include assistance with assessment, and that teachers can check their own teaching against the national outcomes. In relation to the subject of Swedish, Lunneblad and Asplund Carlsson (2010) have shown how national tests in year 5 served to shift the focus from learning processes to performance and measurable results.

Teaching to the test is a particular form of washback, in which teaching focuses on strategies for passing the test and students doing practice tasks resembling those of the test (Bailey, 1996; Menken, 2006). However, research has been unable to prove that this type of guidance before a test actually helps students to improve their results (Bailey, 1996). Popham (2001) distinguishes between 'item-teaching', representative of negative washback and 'curriculum-teaching', which represents positive washback. Teachers engaged in item-teaching design their teaching to practise authentic and devised test tasks, while curriculum-teaching is aimed at specific areas of subject content or skills needed to do the test tasks (Popham, 2001).

Policy enactment

Focusing on how policy is implemented in schools is a way of trying to understand how political governance can be translated into functioning classroom practices. Enacting policy always involves more than just implementation, since policy comprises complex processes linking classrooms to abstract political priorities, according to Ball *et al.* (2012). Different policies convey different ideas of what is desirable, or not, but policies only rarely state explicitly how something is to be done, which means that those who enact policies have considerable scope for interpretation. In order to obtain an understanding of the implications of policy changes for teachers, teachers' double roles as policy subjects and policy actors must be examined (Sikes, 1992; Ball *et al.*, 2012). Swedish teachers' choices and interpretations with regard to the national test in the subject Swedish can be seen as policy actions in which they play the double role of being policy subjects, who have to implement a national test they have had no say in, while being policy actors, who must contribute to enacting intended policies in their classrooms. According to Ball and Olmedo (2013), some teachers respond to this pressure by resisting dominant discourses. Mickwitz (2015) has, for instance, shown how teachers formulate a counter discourse emphasising their professional autonomy against the demand for measurable results. Lundahl and Tveit (2014) argue that the current focus on formative assessment may be regarded as teachers' resistance to external policy governance.

Here we focus specifically on how teachers navigate policy in the form of national tests. The Swedish government took the decision to implement national tests and tasked certain university departments with designing them. Several processes of interpretation have therefore already taken place before students write the national tests. Previous international studies indicate that policymakers overestimate the capacity of introduced external tests to influence teaching on the premise that all teachers interpret tests in the same way and all teachers prioritise the tests in their classrooms (Grant, 2001; Cimbricz, 2002). Gilmore (2002) has nevertheless shown that collective assessment has a positive impact on teachers' assessment practices and professional development, because it allows teachers to develop a vocabulary for discussing students' knowledge and skills. One way of investigating how teachers enact policy in their everyday classroom practices is by studying the washback of the test on Swedish teachers' reflections on the choices they make when teaching.

Another important aspect of enacting policy and the washback that can be seen in relation to the national test in Swedish, is what Ball (2003, 2006) calls performative pressure. The concept can be described as a technique used to govern different operations by constantly holding actors accountable by demanding results that can be reviewed and compared. In schools, this has meant a direction towards a controlling system in which the focus is no longer on visions and ideals but on methods and techniques for finding the most effective ways of using resources and thereby improving results (Marshall, 1999; Perryman, 2009).

The performative society constantly demands improved results, which may lead teachers to adapt their teaching consciously, and notwithstanding their own convictions, in order to fulfil performative requirements and for external inspection. This can mean that teachers at the same time demonstrate what Ball (2003) calls

resistance and capitulation when discussing how some teachers react to performative pressure (so-called fabrications). According to Ball, this can become problematic when teachers face the dilemma of being compelled to prioritise between what can be seen as the school's best interests and what they regard as the best interests of their students. The development of teaching practices risks being driven by the requirement to improve results, something that has also been shown in a recent Swedish study (Vlachos, 2019; see also Hedman, 2019). Changes to educational policy are therefore not only the means of effecting change in the educational system, but are also governance technologies intended to change teachers and their teaching. As Perryman (2009, p. 616) argues: 'It is through the increasing culture of performativity and accountability that conformity, discipline and normalisation is ensured'. Thus, by studying how teachers of Swedish at upper secondary school in Sweden position themselves in this performative landscape, we may understand more about how external standardised tests affect classroom practices.

Method and material

The material, on which this study is based, was collected as part of a larger study involving qualitative semi-structured interviews with teachers of Swedish as mother tongue during the school year 2015/2016. The focus was on teachers' understanding of reading instruction in the subject of Swedish in a performative context Larsson (2018). As mentioned above, the course in Swedish that all students take during their first year of upper secondary school, comprises a national test that consists of three subtests of students' written and oral skills and their reading comprehension. For each interview session, some common questions on different topics were prepared as well as individual follow-up questions. During the interviews the teachers were asked to describe how they selected content for their reading instruction and the choices they made with regards to different texts and students' assignments. They were also asked to describe their perception of the aim of reading within the subject of Swedish, how to assess students' reading competence and the challenges of reading instruction in the context of the current curriculum.

Despite the fact that the study focused on reading instruction in the first year of upper secondary school and the underlying principles determining teachers' text choices and what they wanted students to learn when working with these texts, the national test featured prominently in the teachers' discussions. Questions on the test were therefore added to the later interviews. To begin with, the national test was not initiated as a topic by the interviewer, but already in the first part of the first interview three of the teachers raised the topic when describing that they wanted their students to do more than just pass the national test. In the final group interview, the teachers were therefore asked how they viewed the fact that the national test was so present in their discussions throughout the academic year. In this article we focus on the specific washback effects of the national test that emerged during the study.

The empirical material comprises four individual interviews, each conducted with five upper secondary Swedish teachers (20 interviews) and one group interview (21 interviews in total). Each interview took about an hour, except for the concluding

group interview, which was one and a half hour long. In total, there is about 23 hours of recorded material. At the time, the interviewed teachers worked at four schools in different municipalities. The individual interviews were conducted in August, November, March and May, and the final group interview in June.

Interviewees were selected on the basis of being current teachers of Swedish at upper secondary schools with some years of teaching experience. A second criterion was that they were teaching at least one class of first-year Swedish during the academic year in which the study was done. Since the study focused on reading instruction, the selection was positive, in that teachers who wanted to participate in the study already had a specific interest in reading and reading instruction. Coincidentally, all five teachers taught Swedish on academic programmes that year, although they also had teaching experience of other upper secondary programmes. Written consent to participate in the study was collected from the interviewees prior to the study, and they were also informed that they could discontinue their participation at any time. Most of the interviews were conducted at the teachers' schools. One of the teachers was interviewed elsewhere, because there was no space available to conduct the interview privately.

All the interviews were voice recorded and continually transcribed over the course of the study. All transcriptions resemble written rather than spoken language, as this facilitates reading during analysis and when quotations are presented (Kvale & Brinkmann, 2014). Based on the transcriptions, a systematic content analysis of the teachers' claims and analyses was performed. The transcriptions were carefully read a number of times and to begin with three themes relating to the subtests—writing, speaking and reading—were identified. These themes were then divided into subthemes concerning the teachers' positive and negative perception of the different tests and keywords such as washback, preparation, curriculum-teaching, item-teaching and performative pressure were then used to sort and code the different themes and subthemes in NVivo11 software. Finally, the themes and subthemes were categorised in relation to different aspects of policy enactment with a specific focus on how the teachers expressed adaptation and resistance to the national test's different subtests. Five teachers, one man and four women, participated in the study. They have been assigned the fictitious names of Anna, Bella, Charlotte, David and Ellen.

Teachers on the washback of the test

The participating teachers were ambivalent to the national test in first-year Swedish, but agreed that both they themselves, as Swedish teachers, and the subject of Swedish had been affected by the introduction of several national tests in Swedish when the 2011 upper secondary curriculum (Gy11) came into effect. The teachers emphasised that they nevertheless were positive to the fact that the Swedish National Agency for Education, the SNAE, 'has taken control' by providing joint guidelines and clear frameworks for teaching through policies and the national tests, so that there is, as Charlotte said, 'still a bit of structure in what is taught at schools'. Charlotte did not regard the policy documents for the Swedish subject as too regulatory, and emphasised that she, despite rather strict frames, still had a lot of freedom to choose what

to focus on in the classroom. Anna also emphasised the importance of having a clear mandate and not allowing teachers to do exactly as they please.

The teachers were also of the opinion that one of the advantages of the test is that it provides guidance as to what should be prioritised in a content-heavy subject such as Swedish and that it supports assessment. Charlotte felt that ‘Swedish is a very open subject, so in some way the test determines the content taught in the course’. David said that the national test has advantages:

The designers of the national tests must have considered the link to the knowledge requirements and core content, and must have thought that these are the essential aspects that must be assessed. So this gives us some support too.

Bella noted that the tests ‘ensure that we achieve more equitable schooling’. Charlotte agreed, saying that the tests help ensure that students across the country are not offered completely different courses in Swedish. It may ‘become a bit chaotic if all teachers only follow their own interpretations of what is important for students to know’, Charlotte said. Above all, this effort to maintain a common standard was experienced as something positive, particularly in comparison with the subjects that do not have national tests. Bella said:

If I compare to a teacher who does not have it [test in their subject], grades can vary substantially depending on the teacher and the type of teaching. The structure of the course can also vary a lot then. But we have a kind of clearer guidelines and you know what will be tested so that you can keep fairly... Although the framework is very open, you still have a framework.

While they had this positive attitude, the teachers also felt that they have not allowed their teaching to be much affected by the introduction of a national reading comprehension test in first-year upper secondary Swedish. They maintained, however, that the tests have received much attention and taken up a lot of time, both in the course and when it comes to assessing and grading. If teachers additionally want to prepare their students for the test, it consumes a great deal of time and resources, according to Ellen. She said:

So much time, and of course money, is spent on buying out teachers from the classroom so that they can sit and mark and grade national tests. / . . . / Yes, but why can't we get a lot of time for some other project we think is . . . So, no . . . national tests are all very well, but there shouldn't be so much hype.

Different preparations for the different subtests

Preparing students for a national test may be regarded as a clear example of a teaching activity that teachers and students would not have been engaged in if there had been no test (Alderson & Wall, 1993). Such an activity is therefore a concrete and direct washback effect of a national test. The teachers who participated in this study engaged in different activities connected to the different subtests in the national test in Swedish. There are few indications that the reading comprehension part of the test directly affected their teaching. However, the other two subtests clearly influenced teaching. The teachers described designing one or more tasks during the course that

largely resembled those of the subtest in oral communication. Bella described, for instance, how she collaborated with one of the natural science teachers to find a topic for an oral presentation that students had to deliver in the same way as the oral presentation on the national test. Although students may be nervous about the oral communication test, the teachers felt that they were usually well prepared by the time the test was taken. Bella said that it could be both enjoyable and sometimes very emotional when students do the oral part of the national test.

According to the teachers, however, most preparation is done for the written test. The teachers said that students experience it as the most difficult subtest, and it is particularly formal writing which needs to be practised. This involves using sources correctly and writing different types of texts, such as reports, investigative and argumentative texts. At the different schools where these teachers worked, the Swedish teaching teams had designed tasks that were clearly designed to resemble the written subtest. The teachers generally choose texts that are used as starting point for students' own writing and the tasks are designed to prepare them for the national tests. 'All of us Swedish teachers do this task', said Anna, and continued:

It is linked to the same knowledge requirement and is a way to get to do this twice. Partly to practise, but also in case you fail once . . . It is linked to several things, it is writing, referencing, reporting and then it is linked to . . . yes this with structure and more argumentative texts.

Curriculum-teaching and item-teaching

Oral and written communication tasks in the course are thus designed as more or less exact copies of the national test. These may therefore be understood as examples of item-teaching (Popham, 2001), because they are designed with the national test as an example and aim to ensure that students are familiar with the test format by the time they take it. These preparatory tasks are also used as contexts in which students can demonstrate their skills outside of test situations. In this way, the tasks may be described both as item-teaching and curriculum-teaching simultaneously. It became clear that these can be regarded as curriculum-teaching when the teachers described both the oral and written tests as well as their own pretest tasks as corresponding well to the syllabus for Swedish and the formulations of the knowledge requirements. The results from the oral and written tests directly form the basis of the final course grade, while this is not the case for the results from the reading comprehension test. Charlotte said:

When you sit there with the grading matrices / . . . / it seems very easy to translate—the student had this result on the oral and written parts and you put it in the matrix and tick it. But with the [reading comprehension subtest] it is a question of 'OK, where should I slot this in?'

The reading comprehension subtest also differs from the oral and written subtests in relation to the preparation and tasks done. There is significantly less concrete preparation and the teachers said that it is hard to prepare students for the test. 'We don't know which texts and questions will be chosen', Charlotte said, for instance, and added, 'There is actually no real way to study for this'. The preparation that is done can be described as item-teaching, in which teachers, with reference to the SNAE's

preparation material (Skolverket, 2014), work through the instructions. This means that the teachers, a few weeks ahead of the test and using the Agency's material, show students the test format and the type of questions usually asked, and how the number of lines provided in the answer booklet indicate the length of the answer the test designers expect.

The dominant understanding of the reading comprehension test amongst the teachers is that reading instruction is regarded as curriculum-teaching and therefore as something that is continually done during the course. In the long-term work on reading during the course, all teachers describe good preparation for the reading comprehension test to be reading, discussing and summarising different texts, as well as writing reports on the texts, particularly on non-fiction texts. Yet, the teachers said, this is something that is always included in regular Swedish teaching and not something teachers do specifically because a national reading comprehension test has been introduced. 'I think we prepare for the reading comprehension test the whole year', Anna said. The teachers emphasised reading literature, especially novels and short stories, but also analysing poetry, which invites discussion of what is written between the lines, as good preparation for passing the test. The teachers insisted that these activities were not because of the national test. Particularly the writing of reports was highlighted as a way of obtaining a good idea of students' reading comprehension, because the teachers were of the opinion that one first has to understand a text to be able to write a functioning report on it. They also pointed to the importance of introducing students to several text types in the classroom. Analysis and discussion of different types of non-fiction texts seem primarily to focus on the fact that the students must produce the texts themselves, rather than on learning to read and understand these types of texts. Because the teachers felt that they were constantly working on reading, the reading comprehension test did not have much of an impact on their teaching.

Navigating a performative pressure

Even if the teachers identified some of the positive effects of the introduction of the national test in Swedish in 2011, they missed the freedom and more creative elements they felt to be part of the previous curriculum (Lpf94). Teachers said that in Gy11, Swedish as a subject is more regularised and the focus has shifted from stimulating students' creative sides to assessment and performance. The national test was also seen as governing the course. 'I feel a bit bound by it and I'm not happy that a national test has been introduced in first-year Swedish', said Anna. The teachers felt that much classroom time is generally devoted to the test during the spring term. On the whole, we can establish that the teachers were both critical and positive towards the test, but they nevertheless chose not to resist too much. Anna said, 'I sometimes feel we have too many national tests, I really do think that, but otherwise I think well, . . . it works'. Ellen expressed similar sentiments: 'I see my own role in this, that I am responsible for ensuring that the students have had the opportunity to learn what is required'. Yet, if she could choose, she would not have included the test on the course.

The quotes above are examples of how the participating teachers simultaneously adapt to and resist the national test in first-year Swedish. There is clearly ambivalence towards the test that could be seen when the teachers discussed their teaching in relation to the test and that is also reflected in the washback on teaching that the different subtests seem to show. On the one hand, the teachers felt that the test contributes to more equal teaching and grading by setting joint frames, and they appreciated receiving some assistance with assessment and grading. On the other hand, they opposed being increasingly controlled in the areas of assessment and subject content through the national test. In the quote above, Ellen showed awareness of how her teaching choices affect students' opportunities to perform as well as possible on the test. She therefore chose to prepare her students, despite the fact that she has to enact a policy that she had no voice in. Here her double role as a policy subject and policy actor becomes clear. Anna experienced being caught between roles as an official who has to follow orders and an experienced teacher who wants to do what she thinks will benefit her students the most. David highlighted that even if he does not consciously plan his teaching on the basis of the national test, the course has been affected by the test to such an extent that he would have spent more time on reading and understanding texts if there had been no test. He noted that he instead placed much more emphasis on developing students' formal writing skills—a direct consequence of the written communication subtest. 'You become so afraid', David said, 'that the students will be bad at writing or reading that you practise using old national tests rather than working on larger, overall issues'. David regarded the subject syllabus as the superior policy governing his teaching, and in doing so he centred his interpretation of what should be the core of teaching on curriculum-teaching. He therefore based his teaching on his own interpretation of the syllabus description of what is—or should be—central in the subject of Swedish. In doing so, David resisted the national test, which is relegated to the background in his classroom, because, according to him, it encourages item-teaching. At the same time, however, he admitted that he resorted to item-teaching in focusing so much on preparing his students for the written communication test.

The interviewed teachers adapted most to the oral and written communication subtests and presented most resistance to the reading comprehension subtest. It therefore seems that the teachers' focus concerning the national tests is on preparing students for the subtests in oral and written communication. Preparations for these subtests have a palpable effect in the classroom, since students repeatedly practise the skills tested and because teachers design tasks that are more or less exact copies of those on the national test. None of the teachers who participated in the study seemed to be aware of the contradictions in their resistance to allowing their teaching to be affected by the one subtest, while clearly adapting their teaching to the other two subtests.

Internal and external performative pressures

Our results show that having to handle national tests affects these teachers in different ways. This can be described both as internal and external performative pressures

affecting their capacity to act. An example of an inner pressure is the teachers' experiences of the SNAE's recommendation to use so-called joint assessment of the tests with the aim of ensuring equal assessment of students' performances. Although the teachers think that joint assessment is positive, this is described as a stressful situation which may involve prestige, since each teacher knows that their own students and their performances will be compared to those of other teachers at the same school. David felt that this affected his planning, because no one wants to be the teacher who has 'a class who only got Es' on the written test.

Further examples of inner pressure involve the school management's demands for teacher accountability regarding the relation between national test results and final course grades. David said that management wants 'to have lists' of results, and that it is very important to fill in the grades students received on the national test as well as their final course grades. 'Then there will be discussions about why these do not correspond exactly', said David, who thought that there is no consideration of how well the tests correspond to the policies.¹ 'Such things affect us', he concluded. This can be seen as an expression of what Ball *et al.* (2012) describe as the delivery chain of performative pressure, which involves teachers' having to be able to explain their grading constantly to those who demand explanations. Most of the teachers in the study experienced that results and grades must be delivered in a way that satisfies parents, the school management, the municipality and the media by constantly being able to justify their grading through documenting knowledge requirements and being transparent towards students about how grading is done. This also applies to the national test. All teachers pointed out that they tried to follow the SNAE's directives that national test results should support the setting of grades rather than determine the final course grades,² but when they did so, their professional integrity was questioned by municipality and media representatives. Bella said that she and her colleagues also felt pressured to ensure that final grades correlate with national test results. According to Bella, local education officials 'use these comparative figures the whole time'. She also thought that the teachers who deviated too much from the national tests when awarding grades were questioned and asked to explain this to politicians and management. At the same time, she increasingly resisted this pressure. Bella said, 'It happens more and more that I think no, I'm not going to give in; I am confident in my assessment or in what I think is important in the subject of Swedish, so it doesn't matter if there is a difference'. In this way, Bella exhibited both resistance and adaptation in relation to the external and internal pressures to incorporate the assessment of the national test in her teaching.

Conclusion

When investigating washback effects of external standardised tests the object is to find out whether the test has changed something in teachers' practice. When discussing their teaching, the results in this investigation show that the teachers look upon the different subtests in different ways, and it is clear that the oral and written subtests have changed the teachers' practice whereas the reading comprehension subtest has had little washback. Depending on from which perspective you observe this, the

writing test which has had most washback according to our teachers can be seen as having had both positive and negative washback on the teaching of Swedish. Our study also shows that the teachers' understanding of the test is that it provides guidance to prioritising the core content of the subject as well as how this content should be assessed. The teachers stressed that the test functions as a resource for achieving equity among schools on a national level.

As mentioned above, the participating teachers accorded different importance to the various subtests, which was clear both in relation to how their teaching is adapted to prepare students for the tests, as well as in the significance of test results in setting the final course grade. The results of the subtests in written and oral communication are easily included in the final grade, while teachers were of the opinion that the results of the reading comprehension subtest are difficult to relate to the knowledge requirements in first-year Swedish. It seems that the participating teachers found it more controversial to speak about the measuring and grading of reading than the measuring and grading of writing and speaking skills. This could explain why they were more inclined to keep to curriculum-teaching of reading than of the other two skills. The results also show that teachers have their own agency as policy subjects who enact policy through negotiating between adaptation and resistance in relation to the external and internal performative pressure brought about by the national test and its implementation. In relation to Ball's (2003) terminology where resistance and capitulation can be seen as end points in a continuum of adaptation to the national test, the teachers in our study do not go as far as capitulation but rather adapt in different ways to the three subtests. The teachers' policy enactment therefore involved adaptation as far as the oral and written subtests are concerned, while they presented substantial resistance to assessing some aspects of reading in accordance with the guidelines of the SNAE, which they felt only loosely corresponded to the subject syllabus.

Just as the teachers concerned exhibited both adaptation and resistance to the national test in first-year upper secondary Swedish, there are indications of both positive and negative washback of the test in the teachers' descriptions of their teaching practices. Receiving assistance in prioritising the contents of the course, joint assessment and support when awarding grades may be regarded as positive feedback of the test. A clear negative washback effect of the test can be seen in the teachers reporting spending less time on reading in first-year Swedish since the introduction of the test. They attributed this to the time devoted to the test, both in implementation and assessment, but also to the fact that they felt compelled to spend a lot of time on preparing students, particularly for the written subtest.

The washback of the national test in Swedish and the teachers' enactment of policy may also be considered in the light of the long teaching tradition of the Swedish subject, in which different attitudes to reading, writing and speaking skills can be identified. It is, for instance, clear that teaching literature resonates with established beliefs and traditions of what constitutes relevant knowledge (Lundström, 2007; Persson, 2007; Martinsson, 2018). The aspects of reading that the teachers thought that the test measured can be said to challenge their view of reading as subject content in various ways. Against the backdrop of the requirement to digitalise teaching, Erixon

(2010, 2014; see also Olin-Scheller & Tanner, 2019) using Bernstein's (2000) conceptual framework, has discussed the extent to which teachers are prepared to compromise on subject content in relation to policy regarding new technology. 'The sacred', Erixon claims (2010, p. 156, our translation), 'concerns the specifics of a subject that distinguishes it from other subjects', while 'the profane' subject content is that which teachers, with very little resistance, can adapt to external requirements and new contexts'. As in our study, the teaching of literature emerged as 'sacred' in these earlier studies. There are also similarities in teachers' approaches to the oral and written subtest in our and Erixon's (2010, 2014) and Olin-Scheller and Tanner's (2019) studies, in that the teachers offered very little resistance to adapting their teaching and embedding the external requirements of the subject. As opposed to sacred reading, the oral and written parts of Swedish subject content can therefore be described as profane.

In an environment in which external tests are accorded increasing significance in measuring schools', teachers' and students' performances, there is a risk that teachers, without reflecting on what they are doing and without being fully aware of how much their teaching is influenced by external tests, start teaching to the test because they think they are bound to improve the results of their students and of their schools (Hardy, 2015; Vlachos, 2019). Based on the results of our study, this may be the case in relation to the written and oral subtests. This could be interpreted as positive washback if you only look at the results from a speaking and writing point of view as these skills have been given more time and instruction in the course. However, given the current formulation of the subject syllabus, knowledge requirements and the national test, the reading comprehension test is not in the same position as the other subtests. The results of our study therefore lead to the question of whether the extent of the different subtests' washback effects may make teachers more inclined to teach contents related to speaking and writing skills that are interpreted as being easier to measure and assess, at the cost of reading instruction. At a time when there is mounting performative pressures, and when more weight is given to measurable results, a test introduced with the aim to improve upper secondary students' writing, speaking and reading could be said to have directly counterproductive washback effects on reading.

Disclosure statement

No financial interest or benefit has arisen from this research.

Geolocation information

The data were collected in Sweden.

Data availability statement

There is no data set available with this paper.

NOTES

¹ For example, in the course studied (first-year upper secondary Swedish) there are no specific knowledge requirements for reading comprehension in the course syllabus; yet one of the subtests is about reading comprehension. Thus, the national test of the course does not fully correspond with the course syllabus.

² The study was conducted before the 2018 government decision that national test results should be given more weight in the setting of grades.

References

- Alderson, C. & Wall, D. (1993) Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Andrews, S. (2004) Washback and curriculum innovation, in: L. Cheng & Y. Watanabe (Eds) *Washback in language testing: Research contexts and methods* (Mahwah, NJ: Laurence Erlbaum & Associates), 37–50.
- Bailey, K. (1996) Working for washback: A review of the washback concept in language testing, *Language Testing*, 13(3), 257–279.
- Ball, S. (1993) What is policy? Texts, trajectories and toolboxes, *Discourse: Studies in the Cultural Politics of Education*, 13(2), 10–17.
- Ball, S. (2003) The teacher's soul and the terrors of performativity, *Journal of Education Policy*, 18(2), 215–228.
- Ball, S. (2006) Performativities and fabrications in the education economy: Towards the performative society, in: H. Lauder, P. Brown, J. Dillabough & A. H. Halsey (Eds) *Education, globalization & social change* (Oxford, Oxford University Press), 692–701.
- Ball, S., Maguire, M. & Braun, A. (2012) *How schools do policy—Policy enactments in secondary schools* (Abingdon, Routledge).
- Ball, S. & Olmedo, A. (2013) Care of the self, resistance and subjectivity under neoliberal governmentalities, *Critical Studies in Education*, 54(1), 85–96.
- Bernstein, B. (2000) *Pedagogy, symbolic control and identity: theory, research and critique* (revised edn) (London, Lanham, Rowman and Littlefield).
- Chapman, D. & Snyder, C. W. (2000) Can high stakes national testing improve instruction: Reexamining conventional wisdom, *International Journal of Educational Development*, 20, 457–474.
- Cimbricz, S. (2002) State-mandated testing and teachers' beliefs and practice, *Education Policy Analysis Archives*, 10(2), 1–21.
- Gilmore, A. (2002) Large-scale assessment and teachers' assessment capacity: Learning opportunities for teachers in the National Education Monitoring Project in New Zealand, *Assessment in Education: Principles, Policy & Practice*, 9(3), 343–361.
- Erixon, P. (2010) Svenskämnet i skärmkulturen. [The school subject Swedish in the screen culture], *Tidskrift för litteraturvetenskap, TFL*, 3–4, 153–163.
- Erixon, P. (Ed) (2014) *Skolämnen i digital förändring: En medieekologisk undersökning* [School subjects in digital change: A media-ecological investigation] (Lund, Studentlitteratur).
- Ferman, I. (2004) The washback of an EFL national oral matriculation test to teaching and learning, in: L. Cheng & Y. Watanabe (Eds) *Washback in language testing: Research contexts and methods* (Mahwah, NJ, Laurence Erlbaum & Associates), 191–210.
- Grant, S. G. (2001) An uncertain lever: Exploring the influence of state-level testing in New York State on teaching social studies, *Teachers College Record*, 103(3), 398–426.
- Hardy, I. (2015) A logic of enumeration: The nature and effects of national literacy and numeracy testing in Australia, *Journal of Education Policy*, 30(3), 335–362.
- Hedman, E. (2019, November 13) Jörgen Tholin: Därför är det svårt att motverka glädjebetyg. [Jörgen Tholin: That is why it is difficult to counteract grade inflation]. *Skolvärlden* [The School World]. Available online at: <https://skolvärlden.se/artiklar/jorgen-tholin-darfor-ar-det-svart-att-motverka-gladjebetyg>
- Kvale, S. & Brinkmann, S. (2014) *Den kvalitativa forskningsinterjun* [The qualitative research interview] (3rd edn) (Lund, Studentlitteratur).

- Larsson, M. (2018) *Resultatet av ett prov. Svensklärare om washbackeffekter av nationellt läsförståelseprov på gymnasiet* [The result of a test: Teachers of Swedish about washback effects from a national reading comprehension test in upper secondary school] (Karlstad, Karlstads universitet).
- Lundahl, C. (2009) *Värför nationella prov? Framväxt, dilemman, möjligheter* [Why national tests? Emergence, dilemmas, possibilities] (Lund, Studentlitteratur).
- Lundahl, C. & Tveit, S. (2014) Att legitimera nationella prov i Sverige och i Norge – en fråga om profession och tradition. [Legitimising national tests in Sweden and Norway—A question about profession and tradition], *Pedagogisk forskning i Sverige*, 4–5, 298–323.
- Lundström, S. (2007) *Textens väg: Om förutsättningar för texturval i gymnasieskolans svenskundervisning* [The way of the text. On conditions for selections of texts in upper secondary Swedish teaching] (Umeå, Umeå Universitet).
- Lundström, U. (2017) Att mäta det vi värderar eller värdera det vi kan mäta? Resultatindikatorer som grund för skolval [To measure what we value or value what we can measure? Results indicators as basis for choosing schools], *Utbildning & Demokrati—tidskrift för didaktik och utbildningspolitik*, 26(1), 43–66.
- Lunneblad, J. & Asplund Carlsson, M. (2010) En prövningens tid. Om det nationella provet i svenska i skolår 5 [A testing time. The national test in the school subject Swedish in year 5], *Pedagogisk forskning i Sverige*, 2–3, 81–96.
- Lundqvist, E. & Lidar, M. (2013) Nationella prov i NO och lärares val av undervisningsinnehåll [National tests in the natural sciences and teachers' choice of teaching content]. *Utbildning & Demokrati—tidskrift för didaktik och utbildningspolitik*, 22(3), 85–106.
- Luxia, Q. (2004) Has a high-stakes test produced the intended changes? in: L. Cheng & Y. Watanabe (Eds) *Washback in language testing: Research contexts and methods* (Mahwah, NJ: Laurence Erlbaum & Associates), 171–190.
- Marshall, J. D. (1999) Performativity: Lyotard and Foucault through Searle and Austin, *Studies in Philosophy and Education*, 18, 309–317.
- Martinsson, B. (2018) *Litteratur i skola och samhälle* [Literature in school and society] (Lund, Studentlitteratur).
- Menken, K. (2006) Teaching to the test: How no child left behind impacts language policy, curriculum, and instruction for English language learners, *Bilingual Research Journal: The Journal of the National Association for Bilingual Education*, 30(2), 521–546.
- Mickwitz, L. (2015) Den professionella lärarens möjlighetsvillkor [The professional teacher's conditions for possibilities], *Utbildning & Demokrati—tidskrift för didaktik och utbildningspolitik*, 24(2), 33–54.
- Olin-Scheller, C. & Tanner, M. (2019). "Den heliga" skönlitteraturen och det uppkopplade klassrummet [‘Sacred’ Fiction and the Connected Classroom], in: A. Nordenstam & S. Parmenius-Swärd (Eds.) *Digitalt. SLÅ, Svensklärares årskrift 2018* [Digital. Annual Book of the Teachers of Swedish 2018] (Stockholm, Natur och Kultur), 151–166.
- Palmér, A. (2011) Vad ska bedömas? Gymnasiets nationella prov i svenska och svenska som andraspråk [What should be assessed? Upper secondary school's national test in Swedish and Swedish as a second language], in: M. Ellvin, G. Skar & M. Tengberg (Eds.) *Svenskämnet i förändring? Perspektiv på de nya kurs- och ämnesplanerna, Svensklärares förenings årskrift 2011* [A changing Swedish subject? Perspectives on the new curriculum and syllabi. Annual Book of the Teachers of Swedish 2011] (Stockholm, Natur och Kultur), 143–155.
- Perryman, J. (2009) Inspection and the fabrication of professional and performative processes, *Journal of Education Policy*, 24(5), 611–631.
- Persson, A. (2017) *Lärartillvaro och historieundervisning: Innebörder av ett nytt uppdrag i de mätbara resultatens tid. [History teaching in the age of performativity: Swedish upper primary school teachers' experiences of a new curriculum]* (Umeå: Umeå, universitet).
- Persson, M. (2007) *Värför läsa litteratur? Om litteraturundervisningen efter den kulturella vändningen* [Why read literature? Teaching literature after the cultural turn] (Lund, Studentlitteratur).
- Polesel, J., Rice, S. & Dulfer, N. (2014) The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia, *Journal of Education Policy*, 29(5), 640–657. <https://doi.org/10.1080/02680939.2013.865082>.

- Popham, J. W. (2001) Teaching to the test? *Educational Leadership*, 58(6), 16–20.
- Shohamy, E. (1993) The power of tests: The impact of language tests on teaching and learning, *NFLC Occasional Papers* (Washington, DC, NFLC Occasional Papers).
- Shohamy, E., Donitsa-Schmidt, S. & Ferman, I. (1996) Test impact revisited: Washback effect over time, *Language Testing*, 13(3), 298–317.
- Sikes, P. (1992) Imposed change and the experienced teacher, in: M. Fullan & A. Hargreaves (Eds) *Teacher development and educational change* (London, The Falmer Press), 36–55.
- Skolverket. (2014) Lärarinformation till kursprov i Svenska 1/Svenska som andraspråk 1 VT 2014 [Information to teachers about the national test in the course Swedish 1/Swedish as a second language 1 Spring 2014] (Stockholm, Skolverket).
- Skolverket. (2017, September 19). *Nationella prov i skolan* [National tests in school]. Available online at: <https://www.skolverket.se/bedomning/nationella-prov/alla-nationella-prov-i-skolan>
- Spratt, M. (2005) Washback and the classroom: The implications for teaching and learning of studies of washback from exams, *Language Teaching Research*, 9(1), 5–29.
- Turner, C. (2006) Professionalism and high-stakes tests: Teachers' perspectives when dealing with educational change introduced through provincial exams, *TESL Canada Journal*, 23(2), 54–76.
- Vlachos, J. (2019) Trust-based evaluation in a market-oriented school system, in: M. Dahlstedt & A. Fejes (Eds) *Neoliberalism and market forces in education: Lessons from Sweden* (Abingdon, Routledge), 212–230.
- Wall, D. (2000) The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System*, 28, 499–509.
- Watanabe, Y. (2004) Teacher factors mediating washback, in: L. Cheng & Y. Watanabe (Eds) *Washback in language testing: Research contexts and methods* (Mahwah, NJ, Laurence Erlbaum & Associates), 129–146.
- Wetterstrand, F., Sundhäll, M. & Lundahl, C. (2017) Att få sina bedömningar granskade [Having your assessments inspected] *Utbildning & Demokrati—tidskrift för didaktik och utbildningspolitik*, 26(2), 75–94.