Teachers’ attitudes towards summative testing in England and Sweden

– A comparative study

Jessica Persson
Abstract

The aim of this study is to investigate and compare teachers’ attitudes to the current testing regimes for the 14-19 age cohorts in England and Sweden. A literature review reveals the historical and political contexts to the prevalent testing regimes in the two countries. One important finding is that the central tests in Sweden have clearly defined objectives whereas the tests’ objectives in England are, at best, insufficiently communicated but at worst not clearly thought through. The study also comprises a questionnaire with the purpose of collecting comparable material of teachers’ attitudes in both countries. Despite the small sample, the results clearly highlight significant discrepancies between objectives and achievements on the one hand and between English and Swedish attitudes on the other. Differences arose between the two countries on how well the objectives of tests are being met. English respondents are more positive towards external marking; meanwhile teaching to the test is seen as a lesser problem in Sweden. Many of these discrepancies and differences in attitudes can be explained from historical and cultural differences to education and assessment systems.

Keywords: national test, examination, effects, centralisation, decentralisation
Acknowledgements

I would like to thank my supervisor Satish Patel for his guidance during the process of writing this thesis and his excellent support in administering the on-line questionnaire. I am also grateful to all the respondents to the questionnaire, without whom this thesis would not have been possible. Finally, I would like to thank Ralf for your invaluable support from beginning to end.
# Table of contents

1  INTRODUCTION ........................................................................................................... 5
   1.2  Aims ......................................................................................................................... 6
   1.3  Structure of thesis .................................................................................................. 6

2  BACKGROUND .............................................................................................................. 7
   2.1  Defining key concepts ........................................................................................... 7
       2.1.1  Assessment and testing .................................................................................. 7
       2.1.2  Summative and formative assessment .......................................................... 7
       2.1.3  Effects of testing ............................................................................................ 8
   2.2  National testing and examination – an overview ................................................. 9
   2.3  Education and examinations in England ............................................................. 10
       2.3.1  Overview of the education system ............................................................... 10
       2.3.2  Historical background on education and testing ........................................ 11
       2.3.3  Current reforms ........................................................................................... 12
       2.3.4  Effects of testing ........................................................................................... 13
   2.4  Education and national testing in Sweden ......................................................... 13
       2.4.1  Overview of the education system ............................................................... 13
       2.4.2  Historical background on education and testing ........................................ 14
       2.4.3  Current reforms ........................................................................................... 15
       2.4.4  Effects of testing ........................................................................................... 16
   2.5  Summary of background ..................................................................................... 17

3  METHOD ....................................................................................................................... 18
   3.1  Questionnaire Design ........................................................................................... 18
   3.2  Data Collection and Procedure ........................................................................... 20
   3.3  Selection of respondents ...................................................................................... 20

4  RESULTS AND DISCUSSION .................................................................................... 21
   4.1  Factual results of questionnaire .......................................................................... 21
   4.2  Questionnaire in England and Sweden ............................................................... 21
       4.2.1  Grading assistance ....................................................................................... 21
       4.2.2  Diagnostic function ..................................................................................... 23
       4.2.3  National monitoring .................................................................................... 24
   4.3  Effects of testing ................................................................................................... 25
       4.3.1  Teaching to the test ..................................................................................... 25
       4.3.2  Stress and pressure ..................................................................................... 26
       4.3.3  League tables and test results ..................................................................... 26
       4.3.4  External and internal marking .................................................................... 27
   4.4  Limitations and shortcomings ............................................................................ 28

5  CONCLUSION ................................................................................................................ 29
   5.1  Future research .................................................................................................... 31

6  REFERENCES ............................................................................................................... 32
   6.1  Published references ........................................................................................... 32
   6.2  Electronic references ........................................................................................... 33

APPENDICES .................................................................................................................. 36

Appendix 1 ........................................................................................................................... 36
Appendix 2 ........................................................................................................................... 42
1 Introduction

Education is one of the fundamental building blocks of societies all over the world. During the two or three last decades, several countries have restructured and reformed their educational systems with the purpose of improving schooling. However, reforms of educational systems over the years have been motivated by cultural, economic, political and social reasons differing from country to country (Söderqvist, 2007:10). Testing is an important part of assessment in many countries and is often subject to reform. It is the aim of this thesis to compare the different attitudes to testing in the education systems of Sweden and England. When doing so, allowance will be made for the extent to which such differences arise through the different sociological and political context in the two countries.

Research from the Chartered Institute of Educational Assessors (CIEA) reveals that a majority of UK adults do not agree that their performance in exams reflected their true ability (CIEA, 2008). This and the current reforms of 14-19 education, and the recent disastrous delay on marking 11 and 14 year-olds’ tests (Office of the Qualifications and Examinations Regulator [Ofqual], 2008c), are all part of an ongoing debate on testing and examinations in England.

After the recent (2006) change of government in Sweden, changes to the education system and new directions with respect to national tests have emerged. The debate on testing and assessment in Sweden (Thelén-Wanler, 2008; Scherp, 2008), however, is slightly different to the debate in England. One of the main differences regarding the national testing system is that in Sweden, it is the teachers’ job to mark and grade the centrally set exams of their own pupils, whereas England has implemented a system of external marking. Ongoing reforms and current debates indicate that both countries are still grappling with the respective setups for national testing. Consequently, studies on national testing and examinations have largely been conducted on government initiative, either to promote change or to evaluate education policies. In Sweden, The Swedish National Agency for Education (Skolverket) has the task of achieving goals set by the Swedish government and Parliament through the Education Act (Skolverket, 2008a). Recent studies conducted by Skolverket (Skolverket, 2005; SOU 2007:28; SOU 2008:27) have served as broad overhauls of the education system with the intention of describing current systems and suggesting changes and reforms. In England, the Office of the Qualifications and Examinations Regulator, Ofqual, is the educational body which regulates qualifications, tests and examinations. Ofqual publishes reports on their investigations and work with qualifications along with research in the same area (Ofqual, 2008b).

The working hypothesis of this thesis is that the systems of national tests in England and Sweden are at opposite ends on a scale measuring the subjectivity/objectivity of evaluating pupils’ achievements. Pupils’ achievements in England are mainly evaluated by examinations which are externally marked. Swedish pupils are assessed partly by national tests marked by their own teacher and partly by assessment elements devised by the teachers themselves.

This thesis, focusing on assessments for 14 to 19 year old pupils, will investigate to what extent teachers in England and Sweden believe that objectives of national tests and examinations are indeed met. I will further provide an overview of the educational context and current reforms in the two countries. Despite the two countries’ apparent differences in education systems and approaches to testing in particular, I will highlight how the two countries can benefit from learning from each other.
1.2 Aims

The focus in this study will be on education at upper secondary school level, and in particular age 14-19. The testing required of pupils at this age is comparable between England and Sweden.

In this thesis my aim is to compare testing at three different levels.

- Firstly, I aim to compare reforms historically and currently taking place in England and Sweden with the focus of identifying centralising or decentralising effects and processes through literature review.
- Secondly, I set out to establish whether there is a match between educational bodies' objectives for summative testing and teachers' attitudes towards summative testing in England and Sweden by using a questionnaire.
- Thirdly, I aim to investigate English and Swedish teachers' attitudes towards different effects of testing, through both a questionnaire and literature review.

By analysing literature and research in the area of summative testing along with using a questionnaire to identify teachers’ attitudes on national tests and examinations, I set out to answer the following questions:

1. What reforms to 14-19 education have historically taken place and are currently being implemented in England and Sweden.

2. What type of objectives are examinations and national tests in England and in Sweden based on?

3. What are teachers’ attitudes towards summative testing at GSCE/A level in England and year 9/upper secondary school in Sweden?

As an aside, when interpreting the results from the questionnaires I will keep in mind that the sample of teachers answering the questionnaire is not a large representative sample. Therefore, the conclusions arising from the analysis of the questionnaires should be interpreted with caution.

1.3 Structure of thesis

The remainder of the thesis is structured as follows. The background section begins with definitions of the key concepts which are followed by some general aspects of summative testing. Because the national assessment systems in England and Sweden described in this thesis both operate within a larger educational context, I have included an overview of the educational history, focusing on (i) summative testing, and (ii) the education systems currently operating in both countries. The background also provides (iii) my literature review, which in combination with the questionnaire results answers to my research questions. The background section ends with a summary of the main differences between the two countries, including an identification of centralisation and decentralisation processes in both countries.

In the method section, I describe the design of the questionnaire together with the collection of data and selection of respondents. The questionnaire is used to answer the second and third level of comparison, my second and third aim, by identifying teachers’ attitudes on the objectives and the effect of summative testing.

The two countries are later compared and discussed in the results and discussion section. Conclusions in relation to my discussion and aims are later presented in the concluding section.
2 Background

When it comes to the selection of research on summative testing and reports on governmental objectives with national testing and examinations, I began with the educational bodies in both countries to get an overview of their material, reports and research. I then moved on to searching for peer-reviewed articles and other research conducted in this area, mainly using the Internet as a source. The material has been used to draw overall conclusions with regards to my aim and research questions, as well as, providing a background and educational context with a focus on national testing/examination: its history, effects and debate.

Research by institutions such as the Swedish National Agency for Education (Skolverket) and the English Office of the Qualifications and Examinations Regulator (Ofqual) has informed the discussion in this section which begins with a short description and definition of key concepts later used in the thesis. The following section provides a general overview of national testing and examination and sets out to present what role summative testing plays in society as a whole. The last two sections provide a deeper insight into the history of England’s and Sweden’s education systems with focus on (i) examinations and national tests and (ii) what current reforms are taking place in both countries.

2.1 Defining key concepts

2.1.1 Assessment and testing
School pupils in England and in Sweden are subjected to various tests and assessments. I begin by defining these terms, as used in this paper.

For the purpose of this thesis, assessment is used as an overarching concept that includes different types of evaluation of a pupil’s knowledge and achievement. Brown (2001:384) states that test is a concept within the wider domain of assessment and that a test “is a method of measuring a person’s ability or knowledge in a given domain”. Tests are, furthermore, “almost always identifiable time periods in a curriculum when learners muster all their faculties to offer peak performance, knowing that their responses are being measured and evaluated” (ibid: 401-402).

Testing is one way of assessing pupils’ knowledge and is referred to in this thesis as a clearly defined task, set at a specific point in time in order to measure a pupil’s achievement. One commonly differentiates between formative and summative testing.

2.1.2 Summative and formative assessment
Summative assessment attempts to measure and summarise what pupils have learned and is usually what people refer to when talking about tests. Summative assessment usually occurs at the end of a lesson unit or course (Brown, 2001:402). Summative assessment is also described as a judgement that summarises evidence of achievement up to a final point in time (Taras, 2005:468). Skolverket (2005:19) defines summative assessment, as assessment of learning, and formative assessment, as assessment for learning. Taras (2005:468) argues that all assessment begins with summative assessment (i.e. assessment of goal achievement) and that formative assessment is based on such summative assessment with the addition of feedback used by the learner. The feedback describes the gap between the achieved level and the required standard
together with instructions on how the pupil’s work can be improved. As such, formative evaluation usually occurs when pupils are in the process of “forming” their competence and skill and with the aim of helping them continue that process.

When summative testing is mentioned in this thesis, it is referred to as examinations in an English context and national tests in a Swedish context. I focus mainly on the English GCSE and A level examinations and compare them to the Swedish national tests at year 9 and during upper secondary school. These sets of tests and examinations are comparable as they have a distinct summative and high-stake character. Furthermore, they are coordinated centrally and are taken at approximately the same age.

2.1.3 Effects of testing

In the literature, the concept of washback is used with various meanings. In general terms, it describes how testing influences teaching. The phenomenon is also known in some published work as backwash (Alderson & Wall, 1993:115; Spratt, 2005:8). Spratt (2005:18) concludes that it is not the examinations in themselves which produce various amounts of washback, but rather other factors such as teachers’ attitudes and the stake of the tests. In line with Alderson and Wall (1993:117), I emphasise the importance of understanding the concept of washback. In what follows, I describe the possible effect of national testing and examinations.

Teaching to the test

The term teaching to the test is a form of teaching which aims at preparing pupils for tests and not primarily at improving understanding of the underlying concepts in a subject. The concern about the effects of teaching to the tests has been that pupils spend too much time learning techniques only to help them cope with the specific structure of a test (Firestone, 2004:2; Ward & Shaw, 2002). According to Sturman (2003:262-263), research has indicated that teachers teaching to the test can cause improvement in test scores. Sturman further concludes that research literature often links this type of test preparation with negative effects. The General Teaching Council for England (GTC) state that teachers in England feel that tests lead to a narrowing down of the curriculum and that the current design of the tests encourages teaching to the test (GTC, 2007:4).

External and internal marking

England uses a system of assessment with external markers. During the marking process, samples are checked by more senior markers. One of the aims with this process is to ensure consistency, as if the principal examiner had marked every single test (Qualifications and Curriculum Authority [QCA], 2008b). Roy (1986:15-16) states that the long history of public examinations together with the widespread use of external examinations provided public qualifications, such as GCSEs and A levels, with credibility amongst schools and in society. Generally, peoples’ expectations are that examination boards, who deliver the public qualifications, should control standards of achievement and act with impartiality.

According to Skolverket (2005:8), Sweden separates itself from many other countries due to the fact that the assessment system is based on internal marking. Teachers in Sweden assess and grade their own pupils. The national tests in Sweden act as a grading help for teachers since it provides guidelines and examples of different levels of goal and criteria attainment.
League tables and test results
The 1990s was characterised by an increasing interest from national and local media on pupils’ performances at GSCE and A level. Schools and colleges were ranked in league tables based on examination results which created a competition between schools and colleges for high-achieving 16-19-year-olds. These league tables still play an important part of the public’s perception of schools and colleges (Hodgson & Spours, 2003:15; Sturman, 2003:261). League tables, or performance tables, are recognised by teachers in England as promoting school improvement and providing information for parents. The league tables are, however, viewed by some teachers as having a negative effect on teaching as they create competition between schools, thus undermining cooperation and collaboration. Furthermore, schools are allegedly emphasising tests results and league tables at the expense of pupils’ long-term educational needs (GTC, 2007:3). The General Teaching Council for England has further proposed a system where only a sample of the cohort, a limited number of pupils and schools, is used to monitor national standards.

In Sweden, however, the national test scores are used as a support tool for teachers and not as final examinations. Results from the national tests for compulsory school (year 9) are presented on the internet but attract very little attention and public debate compared to other countries (Skolverket, 2005:26). National tests in upper secondary school are also analysed and presented but only a selection of colleges are chosen (Skolverket, 2008b).

Stress and pressure
Alderson and Wall (1993:118) describe qualitative studies which have shown that the publications of test results have created fear, guilt, shame and embarrassment for teachers. According to Alderson and Wall, some teachers felt that the test results were being used against them. Fear and pressure of performing abnormally on a test, and the consequences of this performance, produce anxiety in pupils; that can be described as being a negative effect of a test. A teacher can similarly fear for poor results by their pupils, which in turn can lead to teaching to the test and a narrowing down of the curriculum. Spratt (2005:9) concludes that testing is described in the literature, as having a narrowing effect on the curriculum. This narrowing effect puts focus on the areas that are most likely to be tested.

In an article from the BBC News (published online on July 14th, 2005), pupils have raised concerns about being under pressure from standards set by society. The pressure is perceived to be very high and greater than pressure placed on past generations. Pupils further emphasise that they are currently tested on their memory retention rather than analytical skill.

2.2 National testing and examination – an overview
Noah and Eckstein (1992:147-170) state that examinations are a means of exerting power over individuals and groups. They note that there appears to be an increasing amount of national debate on aspects of assessments. The choice of assessment instruments and their management are particularly under increased inspection. Eckstein (1996:233-240) concludes that the reason for this development is the increased importance given to assessment results at a personal and at a national level. At a personal level, examination results can largely determine a pupil’s access to further and higher education, and their later social and economic progress. At a national level, examination results have become an indicator of a nation’s economic and social health.

National examinations may also be used as a vehicle to change schools, the education system or society as a whole (Noah and Eckstein, 1992:147-170). Changes in national examinations are often designed to affect instructional content (syllabus) and especially the effort made by pupils.
Changes to examinations are also ways of shifting the control of education either towards, or away from the centre, thus having either centralising or decentralising effects. Eventually examination changes may be used to influence the proportion and socio-economic composition of pupils who gain the opportunity to continue with further education or proceed to desirable job opportunities. The increasing effort from governments in both countries to develop and further characterise vocational education is an example of wanting more pupils to stay on in education and is described by the National Assessment Agency (2008) and in SOU 2008:27.

Traditionally, tests have been used as a selection tool for a pupil’s progression and selection into further education. The concept of “selecting-out” is described by Noah and Eckstein (1992:147-170) as the traditional way of hiring people and granting admission to university education based on a pass or fail of examination. This process of distributing pupils in the educational system was originally designed to allocate pupils on a limited number of study placements and work experiences. Examinations have traditionally been used for “selecting-out” pupils. An alternative use of examinations is “selecting-in” which is a process where the emphasis on passing or failing exams is reduced and examination results are only one element of the counselling and the guidance instruments available. Pupil portfolios and pupils’ profiles may serve as other assessment elements.

Noah and Eckstein (1992:147-170) claim that governments use the results of national tests and examinations as a measure of policy effectiveness, and not just as a means of ranking individual pupils. This role is also highlighted in SOU 2007:28, claiming that it is ever more difficult to design an assessment regime that meets all these diverse objectives. Likewise, the GTC (2007:4) has indicated that national tests and examinations are used for too many purposes: this compromises the reliability and the validity of the tests.

To sum up, the function of national tests and examinations differs in both countries. In Sweden, they function as a relatively small component of the overall grading process. This contrasts to England where they function as being the most important factor when determining grades (SOU 2007:28).

2.3 Education and examinations in England

2.3.1 Overview of the education system

The English education system is divided into primary, secondary, further, and higher education. Compulsory education (primary and secondary school) lasts for 11 years when pupils are between the ages of 5 to 16. This is the most common system although others are in place where parents have the choice to send their children to middle schools and grammar schools: remnants from earlier education systems. In primary and secondary school, pupils pass through stages in education called Key Stage 1-4. Pupils’ achievements are measured through tests at the end of Key Stages 1, 2, and 4.

The focus of this study is mainly on (i) the General Certificate of Education, GSCE, which is normally taken at the age of 16, at the end of Key stage 4, and (ii) the General Certificate of Education Advanced Level (GCE A level), normally taken after the age of 16, upper secondary school (Söderqvist, 2007:67-68). There are currently more than 45 GCSE and 40 GCE A level subjects for pupils to choose from (Ofqual, 2008a). A GCSE or GCE A level is made up of components that can be either externally or internally assessed. Internal assessment, often called coursework, is completed over a longer period of time and is marked by teachers. External assessment is usually conducted through a written examination marked by examiners of an
awarding body (QCA, 2008a). Thus, teachers are not in direct contact with the marked examination papers. Institutions at a school or college get general feedback on each question (J. Alker, personal communication, January 2, 2009).

Post-16 education (upper secondary school) in England is undergoing several reforms (see further description below under “Current Reforms”), but has traditionally consisted of pupils studying for the GCE A Level, commonly known as A levels (henceforth referred to as A levels in this thesis). Pupils study for A levels over a period of two years and usually take on 3-4 subjects each year. Pupils are mainly assessed through examinations in their chosen subjects after each year. Most pupils gain entry to university through their A level qualification (Ofqual, 2008a).

2.3.2 Historical background on education and testing

Education in England, as experienced today, has its origins in the 1944 Education Act, which emphasised the right for children to be educated according to their age, ability and aptitude. It also strived towards creating equal opportunities for all learners, while making allowance for the individual child. At the time of the introduction of the Education Act, schools were responsible for shaping the curriculum and examinations were supposed to reflect the curriculum as opposed to leading it (Roy, 1989:1). The leading role of examinations changed in the beginning of the 1960’s when school based examinations were seen as problematic and schools were accused of being too demanding. During the late 1960’s, this criticism was reversed. School curricula was charged with being unstructured; instead critics argued in favour for an educational system with clear winners and losers and the use of examinations to distribute the prizes (Noah & Eckstein, 1992:147-170).

According to Söderqvist (2007:70), the 1970s was a time characterised by great local freedom. Consequently, schools in England taught more or less the same subjects but with significant local variations such as differences in time devoted to different subjects. Major changes to this local freedom in education took place in the 1980s when the National Curriculum was introduced through the Education Reform Act (1988). Noah and Eckstein (1992:147-170) claim that the changes were introduced to (i) limit schools’ and local authorities’ educational choice, (ii) to announce school-by-school results, and (iii) to subject pupils to regular testing. Söderqvist (2007:70) explains that the new curriculum provided three core subjects: mathematics, science and English, along with seven foundations subjects, including modern foreign language, history, geography, art, technology, music and physical education. However, due to increased workload and lack of time for teachers, further changes to the National curriculum were conducted from the mid 1990s, resulting in the use of both teacher assessment along with a national system of Standard Attainment Tests (SATs) for pupils at the ages of 7, 10, 14 and an examination for “all”, the GCSE (General Certificate of Secondary Education) at the age of 16. These tests had the purpose of comparing schools since teacher assessment was considered less reliable.

One of the major reforms with introducing the GCSE compared to the previous system is the fact that this qualification was intended to be available for the great majority of learners, whereas the previous system had long been criticised for dividing the pupil population (Roy, 1989:42). This division can also be referred to as a traditional gatekeeper function (Noah & Eckstein, 1994) where only a limited number of pupils are certified for completion at satisfactory level of achievement and only a limited number are selected for higher education. This shift, from examinations only being presented for a selected few to being available for “all”, has taken place in many countries.
To summarise the changes made for compulsory school, the Education Reform Act (1988) marked a time when the English central government increased their control over education. This centralisation process featured limitations in local authorities’ educational choice, new testing and assessment arrangements, a reduction in examination boards along with the introduction on the National Curriculum (Noah & Eckstein, 1992:147-170; Söderqvist, 2007:70).

Since the early 1950s, at upper secondary school level, pupils have ended their education in an A level exam, (The Advanced Level General Certificate of Education), designed as an entry qualification for higher education. Since the 1960s, the A level has been graded according to a norm referenced system. This means that only a given percentage of pupils can be awarded a certain grade. A move towards a more criterion referenced grading system took place during the 1980s and was used up until the introduction of the new curriculum in year 2000 (see description below). Currently, A levels are seen as being neither fully norm-referenced nor fully criterion-referenced (Education and Skills Committee, 2003).

Hodgson and Spours (2003:5) argue that the English government has been reluctant to reform upper secondary school or 14-19 education in a coherent way since the mid 1980s. Changes to the A level took place in 2000 and it became known as Curriculum 2000. It was seen as an attempt to steer 14-19 education in a more unified way. One of the major changes in Curriculum 2000 was to split the A level (16 to 18 years) into Advanced Subsidiary (AS) and A2 levels. Pupils are today able to obtain AS levels in a minimum of three subjects during the first year and continue to the full A2 level during their second year. Pupils are encouraged to take a broader range of subject in their first year (e.g. four or five units) and then go on to specialising in these subjects in their final and second year. Pupils can also choose to exit the education system after one year when completing an AS qualification. Hodgson and Spours view the reforms in the Curriculum 2000 as preparations towards a more unified system of learning. Hodgson et al’s (2004:443) analysis suggests that the English curriculum and qualifications’ system have been moving in a unified direction since the late 1970s with both Labour and Conservative governments introducing reforms along the same lines. Another major trend between the late 1970s to early 2003 has been the movement towards a more centralised and unified framework with an Education Ministry: a single regulatory body for curriculum and qualification and a more unified post-16 inspection agency (Hodgson & Spours, 2003:25).

2.3.3 Current reforms

One of the aims of reforms to 14-19 education in England, which are currently taking place, is to increase the participation in education at the age of 17 from 75% to 90% (National Assessment Agency, 2008). The most recent changes to 14-19 education are described in the Green Paper (2008). The Diploma, a new qualification for 14-19 year olds, is set to supplement the current GCSEs and A levels. It has been designed to provide increased practical experience and work-related skills for pupils along with classroom learning, where pupils can further develop their abilities in English, mathematics and ICT (Information and Communication Technology). The purpose of the new qualification is to attract a broader spectrum of pupils. One of the main objectives is to encourage pupils who benefit from studying more practical subjects to continue lifelong learning. It is intended to be flexible in offering both a depth and breadth of study along with providing the skills and knowledge pupils need to go on to employment, college or university. Some schools and colleges across England can currently provide five Diploma courses (Construction and the Built Environment, Creative Media, Engineering, Information Technology, and Society, Health and Development). The intention is to develop a further twelve Diploma courses over the next three years (Directgov, 2008b).
The A levels together with the GCSEs are also undergoing reforms. The new A levels have been taught as from September 2008, and the new GCSEs are to be taught as from 2009. GCSE examinations will have a broader range of question styles with the purpose of allowing all pupils to demonstrate their knowledge and skills. According to the Green Paper (2008:26), changes to assessment at A level will result in more open ended questions on examinations. They are implemented to emphasise deep learning and ability, and to make connections across subjects. The burden of assessment is also meant to be reduced by moving from six assessment units to four in the majority of subjects. Furthermore, a credit-based framework is being developed, affecting both GCSE and A levels, with the purpose of providing a common system for assessing the size (credit value) and the level of different qualifications available for 14-19 year olds. The new framework is designed to create greater flexibility in changing routes of education since it allows for a common understanding when transferring achievements in to a different qualification (Green Paper, 2008:35-39; Department for children, schools and families, 2008).

### 2.3.4 Effects of testing

The following section presents an overview of some of the effects of testing and provides a background to topics discussed in relation to the critique on examinations presented by the Chartered Institute of Educational Assessors (CIEA).

According to the CIEA, pupils in England today sit an average of seventy formal examinations. A recent study by the CIEA revealed that 77% of adults in the UK agreed that their performance on examinations did not reflect their true abilities. Moreover, a poll of CIEA members found that 90% of teachers did not believe that examinations were the best indicator of pupils’ abilities (CIEA, 2008).

Paton (2007b) describes how current examinations are charged with encouraging teaching to the test. It is claimed by an examiner in the article that pupils provide memorised and often identical answers on the examinations. Questions from previous years often reappear on examinations and thus lead to improved results. On the other hand, Paton (2007a) describes how improvements to results and in top grades at GCSE and A level are claimed by Doctor Jim Sinclair (Director of the Joint Council for Qualifications) to be the proof of hard work and dedication from both teachers and pupils. According to the article, the Conservative party charged schools with causing inflation to the league tables by promoting easier subjects and steering pupils away from more difficult subjects. Employers and companies are also expressing their concern over pupils’ skills in English and Mathematics and how well prepared they are for the workplace (Smithers, 2006). Paton (2007b) highlights an interesting viewpoint on the topic by Mary Bousted, general secretary of the Association of Teachers and Lecturers. Bousted claims that the improvements in achievement do not reflect pupils’ actual skills. She charges the national curriculum and assessment system for promoting shallow learning at the expense of a real understanding of a subject. Bousted states that instead of encouraging curious learners, the current curriculum in England is to blame for turning pupils into cynical test-takers.

### 2.4 Education and national testing in Sweden

#### 2.4.1 Overview of the education system

Sweden has a system of a nine-year compulsory school where pupils usually start at the age of seven and finish at the age of fifteen. Pupils can then choose to continue on for three years in upper secondary school. The academic and vocational programmes of study at upper secondary school currently have the same format, where all pupils study for three years. They take the
same core subjects and get a basic eligibility for higher education studies. Since the introduction of a new curriculum (Lp094) in 1994, Sweden has separated itself from other countries regarding its decentralised system with a large local freedom to design education (Söderberg & Wirén, 2004:37-46). SOU 2008:27 (53) describes the lack of an upper secondary diploma as something that further separates Sweden from many countries. This type of qualification often implies the use of external assessors to promote a unified national level of education.

SOU 2008:27 (53) describes the Swedish upper secondary school as being characterised by local freedom. One aspect of this local freedom to design education is expressed in the great variety of study programmes available for pupils. Education at upper secondary level formally comprises 17 programmes of study. However, there are many local varieties of these programmes and a large number of local course are available. This development has been triggered by local authorities. They have created specially designed programmes and courses to attract pupils away from independent schools.

Another aspect of this local freedom is Sweden’s decentralised grading system where teachers have the primary responsibility for judging the performance of their pupils and for awarding grades (Lekholm & Cliffordson, 2008:181).

2.4.2 Historical background on education and testing

Since the mid 1940’s, national assessment has been used in Sweden to develop a fair and reliable grading system. From 1962 and onwards, Sweden has had a 9-year compulsory school. Upper secondary school also originated from around this time. In 1972, a new system was adopted and since then no single national examination (studentexamen) when leaving school has been required for admission to university (Skolverket, 2005:6). The change from high-stake, national examinations was partly motivated by reducing exam stress on pupils. The new system instead emphasised marks gathered through long term course work and nationally set tests. The gains with the new system were judged to be as important as the high degree of comparability achieved by previous national examinations (Noah and Eckstein, 1993). The changes during the 1970s also concerned grades which were administered locally according to national guidelines and set only by the teachers both in compulsory school and upper secondary school. Around this time, the curricula and the schools system were highly centralised and teachers were limited to specific instructions on what and how to teach. The national tests that were introduced in the 9-year compulsory school and upper secondary school were not supposed to determine individual grades for pupils. The function of these national tests was rather to provide teachers with a benchmark: a mean grade score, indicating the class’ average achievement (Noah and Eckstein, 1992:147-170; Skolverket, 2005:7).

Since the mid 1970s authority over schools has shifted from the centre to local authorities (Noah and Eckstein, 1993). Skolverket (2005:11) and Söderqvist (2007:47) describe a continued decentralisation of compulsory and upper secondary education throughout the 1990s. During the reform period in the early 1990s, Sweden implemented a decentralised grading system where grading is primarily based on classroom assessment and criterion-referenced grades. This system is used for providing (i) information on an individual pupil’s acquisition of knowledge, (ii) an evaluation of the educational system, and (iii) an instrument for selection to the next level of education. Lekholm and Cliffordson (2008:186) further explain how the system relies on the underlying assumption that teachers interpret the grading criteria in a similar way and make evaluations that are fair in relation to the syllabus’ goals. However, this decentralised grading system has been criticised for being very subjective. They also describe research which has
indicated that evaluations may differ between teachers, thus creating differences between schools.

One aspect of the decentralised educational system was the introduction of Lpo-94. This reform emphasised a greater local freedom on how to evaluate and how to reach the national goals (Söderqvist, 2007:48). Along with the new curriculum, Lpo-94, came changes to the grading system and the function of the national tests now stated that national tests should:

- contribute to increased goal attainment by pupils
- exemplify course goals and grading criteria (assessment guidelines)
- assist in the process of setting fair and reliable grades (grading assistance)
- show pupils’ strengths and weaknesses (diagnostic function)
- indicate the extent of goal attainment (monitoring) by collating results

(Skolverket, 2005:18)

Furthermore, national tests should not:

- influence the choice of teaching content and teaching methods (since this should be determined by teachers and pupils)
- function as final examination tests (as teachers should award grades based on the assessment of the pupil’s accumulated work or demonstrated knowledge and skills)

(Skolverket, 2005:19)

The Swedish national tests are produced centrally and the content of the tests are confidential beforehand (Lekholm & Cliffordson, 2008:187). National tests in compulsory school are given in Swedish, Swedish as a second language, English and Mathematics. The tests are compulsory for teachers to conduct in the final year of secondary school (9th grade, age 16). They can also be taken on a voluntary basis during the 5th grade (age 11); although about 97% of all schools do undertake the tests. The national tests in upper secondary school are made up of the same subjects as for compulsory school. However, the difference is that subjects in upper secondary school are divided into several courses. As individual schools are allowed to organise their curriculum and courses in whatever way they want, national tests are needed at different times during a year or term (Skolverket, 2005:22).

According to SOU 2007:28 (31), national tests in Sweden are used for assessing development of knowledge and as support for grading. The Swedish national tests are claimed to contribute to equivalence in assessing pupils’ knowledge by illustrating and exemplifying levels required for different grades. Teachers in both compulsory and upper secondary school mark their own pupils’ tests and use scoring and assessment guidelines that facilitate the awarding of grades. No external referees are involved in the marking or grading process. The detailed scoring and assessment guidelines provide examples of pupils’ responses from pilot studies. The guidelines discuss and comment on pupil responses in relation to course goals and grading criteria. There is also a diagnostic function with the tests since the teachers analyse their own pupils’ responses (Lekholm & Cliffordson, 2008:186-187; Skolverket, 2005:22).

2.4.3 Current reforms

It has been the Upper Secondary Education Reform Commission’s assignment to propose a new structure for education at upper secondary school in Sweden. The Ministry of Education and Research will draw up a government bill from the inquiry’s proposals and the Swedish Parliament is due to make a decision in the spring of 2009 (Government Offices of Sweden, 2008). Some of the propositions made in the inquiry SOU 2008:27, which will affect the
national testing system, concern making national tests compulsory in more courses. One of the aims is to give other subjects a higher value. This change is expected to be both expensive and difficult to administer. The inquiry’s proposal is therefore to make the tests more open, not demanding the current secrecy, and making the tests available during the entire course of study, not only as the current end-of-course tests. According to the Government Offices of Sweden (2008), other main features of the reform include changing the number of programmes of study from the current 17 to 19. Out of these 19 programmes of study, five will be preparatory programmes. These programmes provide a qualification that will prepare pupils for higher education, which include Economics, Arts and Humanities, Social Sciences and Media, Natural Sciences, and Technology. Another 14 will be vocational, thus leading to a vocational qualification such as Vehicles and Transport, Health and Social Care, Hotel and Tourism, and Aviation Engineering. Pupils on vocational programmes can make active choices during their course of study which will lead to basic eligibility for higher education. Furthermore, a vocational programme can either be studied in school or as an apprenticeship. Whatever the type of study (vocational or preparatory), pupils will all be studying the same core subjects: Swedish or Swedish as a second language, English, mathematics, history, religious studies, social studies, natural science, and sport & health.

2.4.4 Effects of testing

Skolverket (2005:8) claims that Sweden separates itself from many other countries with regard to the marking process of the national tests. Although cooperation between teachers is recommended, teachers mark the tests of their own pupils. The advantages to this internal marking process are that (i) teachers feel they have respect and trust from authorities, and (ii) it is a competence enhancing process for the teachers. On the other hand, the internal marking process has the disadvantage of creating a greater workload on teachers. The main criticism from teachers is that marking the tests consumes too much time. However, different surveys have shown that teachers view national tests as important assessment tools and express a high confidence in the system. Some teachers think that the national tests are better suited for academic programmes of study, which prepare pupils for higher education. According to SOU 2008:27 (78) views are expressed that the tests should be taken in more subjects.

The concept of teachers marking their own class is however, still supported by a majority of Swedish teachers. A risk with this type of internal marking system is that teachers, compared to external markers, can be more tempted to “mark-up” their own pupils. The informal consensus, however, is that the advantages of the decentralised internal marking system outweigh the disadvantages.

The functions of the Swedish national tests are described by Skolverket as being both formative and summative; as a consequence, negative aspects of high stake tests, such as teaching to the tests, are avoided. The national tests and the national test scores are only to be used as a support tools for teachers and not as final examinations. This supporting function is further emphasised by the fact that it is compulsory for the teachers to conduct national tests but not for the pupils to sit the tests, even though the majority of pupils do this. Pupils’ results on the national tests for compulsory school are presented on the internet and grades are analysed by gender and ethnic background. Compared to many other countries, the publication of results at a school level has attracted very little attention and public debate in Sweden (Skolverket, 2005:26).

SOU 2007:28 is an overhaul of the compulsory school system, issued by the Swedish government in 2006, which concluded that the current curriculum is unclear and vague. This has contributed to different local interpretations of the curriculum. Local syllabi and grading
criteria creating large differences in knowledge requirements have questioned the equivalence of education across the country. Leif Danielsson, the investigator behind the report, concluded that, while norms and values from the latest curriculum (Lpo94) have been implemented, it is the view of knowledge emphasised in previous curricula that has affected the contents of teaching to a high degree (ibid: 188).

Currently, national tests should only constitute one of several ways of assessing pupils’ knowledge, but at the same time, on a national level, the distribution of pupils’ scores for the national tests should roughly correspond to the distribution of grades actually set by the teachers. In his report, Danielsson proposes a change to three occasions of national assessment at year 3, 6 and 9. The national tests taken on these occasions are suggested to have a larger impact on the final grades than before. The report suggests that such a change would ensure a more equal educational experience across the country. One of Danielsson’s suggestions to ensure an improved quality in assessing knowledge is to let two teachers co-operate when marking national tests, a process already implemented in many schools on a voluntary basis.

### 2.5 Summary of background

England has seen a centralisation process in education, shifting power from local to central authorities. Current reforms to education in England continue this trend, attempting to introduce common qualifications framework. Sweden, on the other hand, has become more decentralised, most notably with the introduction of the new curriculum (Lpo94) in 1994.

Looking more closely at the testing system itself, the two countries implement two very different marking systems. England uses external assessors and markers. Meanwhile in Sweden, teachers mark and grade their own pupils. In England summative tests have a greater influence on a pupil’s final grade as compared to Sweden. In both countries centralised tests have spurred the development of league tables, which are given much greater importance in England.
3 Method

This section deals with research methodological issues. In the next subsection I detail issues arising in the design of the questionnaire. Subsection 3.2 is a short description of administrative and procedural issues. The section then concludes in 3.3 with an overview of the selection of respondents to the questionnaire. The complete versions of the questionnaire can be found in Appendices 1 and 2.

3.1 Questionnaire Design

To answer my first research question on teachers’ attitudes towards summative testing, I constructed a questionnaire (see appendix 1 and 2).

Brace (2004:13-14) provides a detailed account on questionnaire design and lists a few problems that can arise within a questionnaire itself.

1. ambiguity in the question
2. order effects between questions
3. order effects within a question
4. inadequate response codes

There are also problems outside the direct control of the researcher.

1. failure of the respondent to understand the question
2. failure of the questionnaire to record the reply accurately or completely
3. asking respondents to describe attitudes on subjects for which they hold no conscious attitude
4. respondents not willing to admit their attitudes or behaviour either consciously or subconsciously

When designing questionnaires, the above issues have to be kept in mind and attempts should be made to minimise their impact. If the respondent fails to understand the question, this could be due to the fact that the respondent does not have the competence to answer it (ibid: 15). Knowledge background was a real issue in the case of my research as teachers in both England and Sweden probably lack experience of each other’s system of external and internal marking. I therefore decided to design different questionnaires, covering the same research questions.

When considering the order in which topics are presented in a questionnaire, it is better to start with the most general topics and questions, and then move on to more specific questions (ibid: 49). I used this structure when designing my questionnaire. The introductory questions were purely factual questions on the respondents’ subject(s) and level of teaching, yet still relevant for my survey. The later questions were focused on the objectives of summative testing and enquire the respondents’ general perceptions. Towards the end, I returned to the issue of the objectives but presented the respondents with more detailed questions such as the possible effects of testing.

Brace (2004:16) points out that questionnaires can be designed in a way that creates failure to record the answers accurately or completely. What I needed to consider when creating multiple choice questions was to make sure there were enough possible answers for the respondents to choose from. I encountered a problem with this as I wanted to know the teachers’ attitudes concerning which point in time they thought was suitable for subjecting pupils to national tests and examinations. As I provided too many alternatives to chose from, thereby creating ambiguous answers, no such results about this are discussed in this thesis. A simple solution to
this problem could have been to make this question into an open question, thus asking the respondents at what age they thought national tests and examinations should start.

When constructing the actual questions, I mainly used closed questions. Brace describes closed questions as usually restricting the respondents’ answer to a yes or no. This type of question is also used when the respondent is asked to choose from a number of alternative answers. Closed questions with more than one answer are also known as multiple choice questions. In contrast, open questions allow respondents to answer in their own words and try to seek a spontaneous response. I used open questions on two occasions. Firstly, in the beginning when the respondents were asked which subject(s) they mainly taught. The use of a multiple choice question here would have been impossible to administer due to the range of teachers possibly answering the questionnaire. Secondly, an open ended question was used to supplement a question on assessment methods (question 9, appendix 1). After asking whether teachers preferred alternative assessment methods to national tests, the open question was designed to give an opportunity for a nuanced expression of preferences.

When designing the multiple choice questions I had a few general rules in mind as described by Brace (2004:67). The most common method of measuring attitudes is an ordinal rating scale\(^1\) where respondents are asked to give an answer using a range of options.

The use of five answer options is common in questionnaires and surveys as it is easy to understand for the respondents. To avoid biased answers from the respondents, it is important to consider the balance of the scale. If a scale has answers such as excellent, very good, good, average and poor, it has a surplus of positive answers which leads to an advantage for the positive answers. If you know, however, that the responses are going to be directed towards one end of the scale, it might be worth having more of those answer options. As the questionnaire writer, I had to consider whether or not to provide Don’t know boxes to the respondents. Brace states that it is in the questionnaire writer’s interest not to encourage Don’t know answers. At the same time it could be important when interpreting the results of a survey to distinguish between respondents who did/could not answer and respondents who genuinely chose the more neutral mid alternative on the scale. The use of an even number of answer options eliminates the more neutral mid alternative and “forces” respondents to take a stand one way or the other. I chose however to use 5 answer options, ranging from Strongly agree, Agree, Neutral, Disagree, to Strongly Disagree, with the addition of a Don’t know box.

My third research question was aimed at answering what type of objectives national tests and examinations are based on. My previous knowledge about the Swedish national tests led me to believe that the English examinations were equally based on similar objectives. However, after making enquiries into the matter I drew the conclusion that English examinations were not based on objectives that were comparable to the Swedish ones. I therefore chose only to use the Swedish objectives when constructing the questionnaire. For the English version of the questionnaire, I chose to include only three of the five Swedish objectives. Since English teachers do not do any marking of examinations, the objective of exemplifying course goals and criteria did not seem applicable to the English version. Neither did the objective of contributing to increased goal attainment by pupils. A system where pupils strive to attain goals appears to be limited to the Swedish educational context.

---

\(^1\) A scale where you can order the alternatives but cannot say “how far” they are apart, e.g. tiny/small/big/huge (Selvanathan A., Selvanathan S., Keller G. and Warrack B., 2000.)
3.2 Data Collection and Procedure

In this section I describe the practical collection of my material and the reasons behind the construction and administration of the questionnaires.

Web-based questionnaires can present questions in the order that the questionnaire writer intended to and, as an additional advantage, web-based questionnaires are easier to administer and faster to analyse (Brace, 2004:40). I chose to design the questionnaire as a web-based, or online, questionnaire and distribute it as a link via e-mail. My main intention with this choice was to generate as many answers as possible while having the benefit of a faster analysis from using a specific program to create graphs and present results.

With each e-mail sent out there was an introduction letter (see Appendix 1 and 2) attached and the same information was repeated at the beginning of the online questionnaire. This letter was written in English for the English version of the questionnaire and in Swedish for the Swedish version. It also offered a justification for questions being written in English for the Swedish questionnaire. Johansson and Svedner (2006:33) point out important information to include in an introduction letter. This led me to include information about myself and state the purpose with the questionnaire. Finally, insurance was provided that their answers would be treated anonymously and information about where the respondents could turn in case they had questions.

3.3 Selection of respondents

The questionnaire was originally sent out to teachers working in an upper secondary school in England and in Sweden. The schools were chosen based on the contacts I had within these schools. Since the number of Swedish respondents was very low compared to the number of respondents on the English version of the questionnaire, I chose to use alternative ways of finding Swedish teachers to answer my questionnaire. I distributed the link to my questionnaire via a forum (with about 780 members) for teachers and teacher pupils on the Internet site Facebook.com (Facebook, 2008). Altogether 114 teachers were sent the English version of the questionnaire, with 55 responding. The Swedish version of the questionnaire was originally sent out to considerably fewer teachers (seven) due to limitations in my access to possible respondents. These seven teachers were further encouraged to forward the link to colleagues. It is difficult to estimate how many members of the Facebook group did indeed see the Swedish questionnaire. I estimate that about 100 teachers were aware of the questionnaire altogether in Sweden with 16 responding. Both questionnaires were sent out a few days before the Christmas break which probably contributed to the low number of respondents.

Due to the fact that the participants were encouraged to send the link to the questionnaire along to other teachers, I have not been able to ascertain the selection of respondents. This was an intentional choice mainly designed to obtain as many answers to the questionnaire as possible. I need to make clear that it is impossible to argue that the feedback received is representative for any larger cohort of teachers. Nevertheless, as demonstrated below, the results do deliver some valuable insights.
4 Results and Discussion

The full range of responses and results from the questionnaire are presented in Appendix 1 and 2. Written comments by respondents are, however, not included in the appendices. In this section, I do not go into detail for each question posed in the questionnaire. The purpose of this section is to highlight interesting results followed by discussion and analysis in relation to my research questions and aims. In hindsight, I realise that some questions in the questionnaire were related to more than one aim. I therefore chose not to present results in a factual manner and isolated but rather in combination with the discussion and research. To begin with in this section, I present the results of the first couple of factual questions in the questionnaire. In section 4.2 which answers my second aim, I present and discuss the results of the questionnaires in relation to previous research, keeping in mind the objectives with national testing as a starting point. In section 4.3, I focus mainly on the results of the questionnaire in relation to the effects of testing as I set out to answer my third aim.

4.1 Factual results of questionnaire

All of the respondents to the English version of the questionnaire claimed to work in upper secondary school. Overall, the English respondents taught a wide range of subjects: 19 in total. About half of the respondents taught English Language and Literature, Business, Maths or Biology, while the remaining half were spread out over 15 other subjects.

From the Swedish respondents, twelve worked in upper secondary school, while four mainly worked in secondary school. The majority of the respondents taught English and Swedish. These are subjects where national tests occur both at secondary and upper secondary school.

4.2 Questionnaire in England and Sweden

Teachers in England and Sweden were initially asked to give their opinion on the different objectives of externally marked examinations and national tests. The following section analyses and discusses teacher’s responses. I also provide possible reasons that affected the respondents expressed attitudes and relate the findings in the questionnaire to previous research.

4.2.1 Grading assistance

One of the goals with national tests in Sweden is to assist in the process of setting fair and reliable grades. In the questionnaires, teachers from both countries were asked to give their opinion on this objective which deals with grading assistance.

From the English teachers’ responses to question 3 and 4 (Appendix 1), it is apparent that all of the respondents see a role for external examination in the grading process and confirm that external examination should continue to assist fairness and reliability in grading. However, only 62% of respondents believe that external examination do, in fact, assist in setting fair and reliable grades. The majority of the remaining teachers (34%) did not think that external examinations currently provide fair and reliable grading assistance, while still seeing a need for it. This group of teachers is clearly unhappy with the current system. I believe one of the contributing factors to their expressed unhappiness could be explained using the results in Table 1.
47% of the respondents do not think that external marking is consistent across the pupil cohort (disagree and strongly disagree in question 11.1 in Table 1). It therefore appears as if an unsatisfactory marking regime is negatively affecting the perceived effectiveness of the system of external examination as a whole. However, my initial hypothesis was that teachers in England would be more hesitant and negative towards the fairness and reliability with external examination due to recent marking delays (Ofqual, 2008c). I think that the strong belief in external examination amongst the respondents, partly, is due to the fact that upper secondary school has not been affected by the great marking delays to the same extent as secondary school (End of Key Stage 2 and 3 tests). The respondents to the English version of the questionnaire all taught at an upper secondary level. Furthermore, the respondents’ faith in external examination acting as grading assistance could also be due to the fact that in England they have traditionally been seen to provide impartial judgement (Roy, 1986:16).

The majority of respondents to the Swedish version of the questionnaire (Appendix 2) thought that national tests currently provide them with grading assistance (question 3.4). As the objective of providing grading assistance has been emphasised since the introduction of the national tests (Skolverket, 2005), this response seems to reflect that the introduction of the national test has achieved this objective; furthermore, this objective has been communicated well to teachers. This is confirmed by Skolverket (2005:29) in several surveys which indicate a high belief in national tests as an important assessment tool among teachers.

The objective of national tests providing grading assistance is of a slightly different nature for teachers in Sweden, since it is the teachers alone who do the marking of the tests. The national tests thus act as important tools for defining course objectives and grading criteria not applicable to the English examinations to the same extent. It is, therefore, impossible to directly compare attitudes of English and Swedish teacher’s with respect to the grading assistance function of summative tests. When reading the above results it has to be kept in mind that summative tests in the two countries have different functions and that teachers are exposed to different testing traditions.
4.2.2 Diagnostic function

In both versions of the questionnaire, I used another objective from the Swedish national tests. This objective aims to provide a diagnostic function, showing strengths and weaknesses in pupils’ understanding of the subject.

A majority of the respondents to the English version of the questionnaire (87%) thought that external examinations should be used to show pupils’ strengths and weaknesses, while the respondents were clearly divided on the matter of whether or not examinations currently provide this diagnostic function. 49% of the respondents were dissatisfied with how current examinations provide them with information about pupils’ strengths and weaknesses. My understanding of the division in teachers’ attitudes is connected to the concept of teaching to the test (described in section 2.1.3). A majority (89%) of the respondents agreed or strongly agreed that the content of their teaching largely was determined by examinations (question 13.8). I would claim that teachers realise that a pupil’s test result may just as much reflect how they coped with the particular type of summative assessment as it reflects their general understanding of the taught material. If this was the teachers’ interpretation, it would explain their dissatisfaction with the central tests’ diagnostic function.

Moreover, it is difficult for English teachers to view an individual pupil’s marked exam papers. Since marking of examinations in England is conducted by external markers (QCA, 2008a), teachers are only provided with general feedback on each question. As such, they are unable to identify an individual pupil’s strengths and weaknesses (J. Alker, personal communication, January 2, 2009). Diagnostic information about individual pupils is thus unavailable for teachers. This could be one of the contributing factors as to why some of the respondents were unhappy with the diagnostic function of examinations. My analysis above is strengthened by a significant group of respondents in question 13.3 (Appendix 1), who indicate that the tests currently do not highlight an individual pupil’s strengths and weaknesses.

In Sweden, the objective of providing diagnostic information about an individual pupil is achieved since teachers mark their own pupils on national tests (Skolverket, 2005). A majority of the Swedish respondents indicated that national tests currently provide this diagnostic function. The respondents further indicated that national tests are helpful in their assessment of pupils (question 8.4, Appendix 2) and in identifying areas of improvement for their pupils (question 8.3). I think the results of my questionnaire yet again reflect teachers’ overall contentment with the system of national tests in Sweden, also described by Skolverket (2005:29).

The differences in assessment systems between England and Sweden make the comparison of the two countries concerning the diagnostic function of summative tests problematic. English teachers do not obtain individual diagnostic information from examinations since they are marked externally.2 On the other hand, Swedish teachers gain a lot of diagnostic information about individual pupils due to the internal marking of national tests. This fully explains the different responses of teachers in England and Sweden. What is interesting, however, is that respondents from both countries agree that the diagnostic function of summative tests is important.

---

2 In England detailed information can only be obtained when a request for a remark is lodged (J. Alker, personal communication, January 2, 2009). This clearly indicates that there is no emphasis on individual diagnostic function of national tests in England.
4.2.3 National monitoring

National tests in Sweden are designed to provide results that indicate the extent of goal attainment in education throughout the country. This last objective, used in the two questionnaires, was aimed at generating teachers’ opinions on the national monitoring function of summative tests.

Question 7 and 8 of the English version of the questionnaire (Appendix 1) was aimed at demonstrating teachers’ attitudes towards the function of national monitoring of external examinations. A majority of the English respondents (90%) clearly stated that external examinations at GCSE and A level should indicate the extent of pupils’ achievements in England as a whole (question 8). However, the respondents were yet again divided about whether or not external examinations currently accomplish the objective of national monitoring (question 7). I think the unhappiness expressed by a group of the respondents could partly be due to issues regarding some of the effects of external marking. This conclusion is based on the results of the questionnaire where the English respondents were clearly divided on issues such as, if external marking is consistent or not across the pupil cohort (question 11.1) and if the administration of external marking is inefficient or not (question 11.2). Some of the respondents were unhappy with certain aspects of the system of external marking. Unhappiness with these aspects could be a contributing factor as to why a group of the English respondents were discontent with how examinations currently provide national monitoring.

Another issue which may contribute to explaining why a significant proportion of teachers do not believe that the national monitoring function is satisfactorily achieved may be due to the commonly held view that the general trend in improved results is not matched by an underlying improvement of knowledge. For example, it is thought that improved overall GCSE results are mainly due to a move to “simpler” subjects (Smithers, 2006; Paton, 2007a).

Certain unsatisfactory aspects apart, the English respondents seemed happy with the current system of external marking. 87% clearly indicated preference towards external marking as opposed to internal marking of examinations (question 12). This was further expressed when the English respondents were asked if they preferred other methods for assessing pupils’ achievements as an alternative to external marking (question 9). 27% of respondents indicated to be more positive towards alternative assessment, preferring a mix between external and internal assessment. My interpretation of these results is that while a group of English respondents see a need for a different assessment system, where external and internal assessments are given a more equal emphasis, there is basically no suggestion that the current system of examinations should be scrapped altogether.

In Sweden a clear majority of the respondents (Appendix 2) agreed that national tests currently provide national monitoring (question 3.5). This was interesting to note since the respondents were divided on several issues concerning some of the aspects of national tests presented in Table 2.

---

3 See Table 1.
Some of the respondents thought that internal marking of national tests was affected by the marker's knowledge of a pupil's identity (question 6.2 in Table 2). Furthermore, the answers to whether or not internal marking is consistent across the pupil cohort were largely spread out (question 6.1 in Table 2). While these results seem to indicate that there may be slight inconsistencies in marking across a pupil cohort, these difficulties, in the opinion of the respondents, do not negatively impact on the national monitoring function of the examinations. One way to explain this attitude is to argue that teachers seem to think that such inconsistencies are averaged out when looked at nationwide.

### Table 2. Summary statistics for question 6 (Appendix 2).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don't know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal marking of national tests is consistent across the student cohort</td>
<td>7.1% (1)</td>
<td>14.3%  (2)</td>
<td>21.4%   (3)</td>
<td>21.4%   (3)</td>
<td>0.0% (0)</td>
<td>35.7% (5)</td>
<td>14</td>
</tr>
<tr>
<td>Internal marking of national tests is affected by the marker's knowledge of the students' identity</td>
<td>0.0% (0)</td>
<td>42.0% (6)</td>
<td>28.8% (4)</td>
<td>14.3% (2)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>External marking of national tests is consistent across all pupils</td>
<td>14.3% (2)</td>
<td>21.4% (3)</td>
<td>28.8% (4)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>35.7% (5)</td>
<td>14</td>
</tr>
</tbody>
</table>

Some of the respondents thought that internal marking of national tests was affected by the marker's knowledge of a pupil's identity (question 6.2 in Table 2). Furthermore, the answers to whether or not internal marking is consistent across the pupil cohort were largely spread out (question 6.1 in Table 2). While these results seem to indicate that there may be slight inconsistencies in marking across a pupil cohort, these difficulties, in the opinion of the respondents, do not negatively impact on the national monitoring function of the examinations. One way to explain this attitude is to argue that teachers seem to think that such inconsistencies are averaged out when looked at nationwide.

### 4.3 Effects of testing

In this section, I present the results of the questionnaire with regards to the effects of testing and discuss my findings in relation to previous research.

#### 4.3.1 Teaching to the test

When looking more closely at questions in the questionnaire dealing with issues such as teaching to the test, there is a clear difference between respondents in England compared to Sweden.

### Table 3. Summary statistics of question 8.9 (Appendix 2).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don't know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>The content of my teaching is largely determined by the national tests</td>
<td>0.0% (0)</td>
<td>7.1% (1)</td>
<td>21.4%   (3)</td>
<td>50.0% (7)</td>
<td>7.1% (1)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
</tbody>
</table>

### Table 4. Summary statistics of question 13.8 (Appendix 1).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don't know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>The content of my teaching is largely determined by the examinations</td>
<td>51.1% (23)</td>
<td>37.8% (17)</td>
<td>8.9% (4)</td>
<td>2.2% (1)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>45</td>
</tr>
</tbody>
</table>
A large majority of Swedish respondents did not seem to think that national tests largely
determined the content of their teaching (question 8.9 in Table 3), while an overwhelming
majority of the English respondents agreed or strongly agreed to the statement (question 13.8,
Table 4). To me this indicates that the issue of teaching to the test is perceived to be more
prominent in England compared to Sweden. It seems reasonable that teachers in Sweden do not
teach to the test to the same extent as English respondents. I base this assumption on the fact
that the purpose of national tests in Sweden is only to act as support in the grading process and
not as a final, high-stake test. It is further stated in the objectives of Swedish national tests that
they should not influence the content or teaching methods (Skolverket, 2005:19). The English
respondents claimed to be teaching what examination syllabi emphasise. This issue of how
current examinations encourage teaching of the test is emphasised in both public debate (Paton
2007b) and research (General Teaching Council for England, 2007).

4.3.2 Stress and pressure

Two of the questions in the questionnaire dealt with possible pressure exerted onto teachers and
pupils from examinations. Respondents from both countries clearly indicated that examinations
and national tests create a lot of pressure on pupils (question 13.6, appendix 1 & question 8.6,
appendix 2). It was however interesting to note that the Swedish respondents seemed to agree
that national tests put pressure on pupils. My hypothesis was that respondents in Sweden would
indicate pressure being lower on pupils compared to English respondents based on the notion
that the Swedish national tests are not designed as being high-stake tests (Skolverket, 2005). My
assumptions were based on the notion that the change in Sweden from a high-stake, national
examination (studentexamen) into the current system of marks gathered on a long term basis
was partly motivated by reducing stress on pupils (Noah and Eckstein, 1993). I have not,
however, conducted a thorough review of research into the area of stress and pressure exerted
on pupils.

The second question in the questionnaire on this issue dealt with pressure exerted onto
teachers. Both the English and Swedish respondents indicated that examinations put a lot of
pressure on teachers (question 13.5, appendix 1 & question 8.5, appendix 2). The Swedish
questionnaire had an additional question dealing with the marking of national tests (question
8.8)4. The Swedish respondents confirmed the results of other studies by Skolverket (2005:29),
indicating that the great workload and time consuming effort to mark the national tests are the
main criticism from teachers.

4.3.3 League tables and test results

The three last questions of both versions of the questionnaire dealt with issues concerning
league tables. The majority of respondents from both countries answered to know how well their
school or college performs on national tests and examinations (question 13.10, appendix 1 &
question 8.11, appendix 2). A majority of English respondents thought that their school or
college seemed more interested in high test scores than improving overall pupil learning
(question 13.11, appendix 1). Interestingly enough, none of the Swedish respondents reported
the same attitudes (question 8.12, appendix 2). My interpretation of these differing results is
largely based on the concept of league tables. Although results on the national tests in Sweden
are presented on the internet, they have attracted little attention and public debate according to
Skolverket (2005:25-26). Furthermore, my understanding is that the attitudes of the Swedish

4 Since English examinations do not involve internal marking this question was not included in the
English version of the questionnaire.
respondents confirm research done by Skolverket (2005). This is in line with research by Sturman (2003:261), which indicates that league tables and examination results play a big part in society’s perception of schools and colleges. In question 13.9 (appendix 1) a majority of English respondents indicated that examinations should not be used for comparing schools. Judging from the results, it seems likely that the English respondents are discontent with the effect of examination results’ current usage in league tables. The respondents’ answers are in line with research presented by the GTC (2007:19) which indicates that some teachers view league tables as having negative effects on teaching, such as creating competition between schools and promoting teaching to the test.

4.3.4  External and internal marking

The Swedish respondents were split on the question about whether they preferred external marking to internal (question 7, appendix 2). According to Skolverket (2005:29) teachers in Sweden are content with the current system of national testing. However, from the results of the Swedish respondents I conclude that some of them are discontent with the current marking arrangements of the national tests. These findings corroborate previous research claiming that internal marking is the main criticism from teachers of the national testing system in Sweden (ibid: 29).

The Swedish respondents were divided (question 6.1, see Table 2) about whether or not internal marking of national tests is consistent across the pupil cohort (it should be noted that 35% did not express a view). Furthermore, the following question (6.2, see Table 2) also indicated somewhat diverse results on whether or not internal marking is affected by the marker’s knowledge of a pupil. My hypothesis was that the respondents would express a higher discontent with the system of internal marking. Despite research (Skolverket, 2005:29) indicating that internal marking is criticised and charged with overloading teachers with work, half of the respondents answered to prefer internal marking. These results could partly be due to the fact that there were respondents to the Swedish questionnaire who did not teach subjects which have national tests. One of the reasons as to why some respondents answered to prefer internal marking (question 7) could be due the diagnostic function provided by this type of marking system (Skolverket, 2005:22). Issues concerning the workload of marking could be a possible factor explaining why some respondents preferred external marking (ibid: 29). A further reason for the respondents’ preference for external marking could be a reaction of possible discontent with the internal marking system. This discontent could be based on the fact that the internal marking system relies on teachers interpreting the grading criteria in a similar way (Lekholm & Cliffordson, 2008:186). This is also indicated in SOU 2007:28 (187) where local interpretation of grading criteria is thought to cause large differences in knowledge requirements.

I intentionally chose to ask respondents their overall preference for external or internal marking after posing questions on more detailed effects of internal and external marking. The placement of the question on overall preference makes it likely that respondents considered pros and cons with both marking systems before answering. Clearly, as discussed before, the respondents found a disadvantage with internal marking, reflected in the response to question 7 (Appendix 2).

Looking more closely at English respondents’ attitudes towards external and internal marking, there was a clear majority who preferred external marking (question 12, Appendix 1). Interestingly enough, there was a clear division between English respondents as to whether or not external marking is consistent across the pupil cohort in question 11.1 (see Table 1). These
results can be interpreted as being contradictive but my understanding is that English respondents rejected internal marking partly because of the extra workload that this marking regime would generate. The respondents however indicated a need for examinations providing a diagnostic function (question 6, Appendix 1). In my view, this function would only be delivered through a system of internal marking.

One of the major differences between the English and the Swedish testing systems is the use of external and internal marking. My conclusion on this issue is that the respondents might prefer external marking because of the decrease in workload. A perceived inconsistency with internal marking (Lekholm & Cliffordson, 2008; SOU 2007:28) could be another reason as to why some of the respondents preferred external marking. The English respondents indicate however that a system with rigorously controlled external markers does not automatically generate consistency in marking across the pupil cohort.

4.4 Limitations and shortcomings

It is a limitation that I was not able to identify clearly defined objectives with English examinations. The approach used translated the objectives for the Swedish national tests as much as possible. The objectives chosen for the questionnaire (grading assistance, diagnostic function and national monitoring) are all taken from the objectives with the Swedish national tests, due to a lack of documentation for similar comparable objectives with examinations in England. Only the two latter objectives could possibly provide a similar translation to the English system, because grading assistance is an important part of the Swedish system, and because teachers do all the marking themselves. The very controlled and rigorous marking system in England does not create a need for grading assistance for teachers to the same extent as in Sweden.

Another shortcoming was that the Swedish version of the questionnaire had a flaw. It did not ask the Swedish respondents whether the national tests should be used to fulfil the mentioned objectives, as the English version did in questions 4, 6 and 8 (appendix 1).

Johansson and Svedner (2006:30) describe the construction of questions as being one of the weaknesses with using questionnaires as a method in thesis writing. Furthermore, Brace (2004:16) mentions ambiguity as a problem when constructing questionnaires, where the respondent (i) cannot understand what is being asked or (ii) understands the question differently from the way it was intended to be understood. I encountered these aspects in writing my thesis. Some of the questions in the questionnaire are formulated in a way that possibly could create different interpretations with the respondents. This ambiguity has further led to difficulties when analysing the results and drawing conclusions. One example is how I intentionally used pupils in plural throughout the questionnaires to avoid potential confusion. This use, however, led to some of the questions not pointing out what I wanted to know. Since Swedish teachers mark their own pupils, I judged it as being apparent to the respondents what was intended in the questions regarding diagnostic function. However, I do not think that the respondents to the English version of the questionnaire fully understood my intentions with the question because of my poor formulation as the question was formulated with the use of the plural pupils. I should have more clearly indicated that the question was aimed at diagnostic function relating to an individual pupil and not with regards to the pupil cohort as a whole or within a class. The problem with ambiguity could have be minimized in my questionnaire by letting a reference group answer the questionnaire, as a means of finding out whether these tests-respondents has understood the questions correctly, before the final questionnaire was
handed out to my “real” respondents. In the short time frame given for this thesis, this was not feasible.

Another weakness, as described by Johansson and Svedner (2006:30), was the difficulty of administering the questionnaires. The number of respondents to the English version of the questionnaire was clearly larger than the Swedish version. My intention was to have a more or less equal number of respondents from each country. Due to the very low number of respondents, 71 altogether, I think my survey has more of a qualitative character, and, furthermore, it is not aimed at describing anything else than a group of teachers’ attitudes towards summative testing.

Originally, it was my intention to supplement the questionnaires with a few more in-depth interviews. This idea only remained an idea due to lack of time, but could have been used as a complement and made the analysis of the results more interesting.

5 Conclusion

My intention with this thesis was to make a comparison at three different levels. Firstly, I began with a comparison of historical and current reforms in both countries, trying to identify centralising or decentralising effects and processes through literature review. Secondly, I set out to establish whether there was a match between educational bodies’ objectives for summative testing and teachers’ attitudes towards summative testing by using a questionnaire. Thirdly, I investigated English and Swedish teachers’ attitudes towards different effects of testing, through both a questionnaire and literature review. The following paragraphs offer conclusions with respect to each of these aims.

With regards to the first aim, research concludes that there has been a shift in authority over education in England where local authority over schools has decreased and has been transferred to the central government. The introduction of the National Curriculum, Curriculum 2000, and a more unified framework for qualifications are manifestations of this centralisation process (Hodgson & Spours, 2003:25; Noah and Eckstein, 1992:147-170; Söderqvist, 2007:70). In Sweden, however, from a very centralised education system prior to the 1970s, the trend has been to transfer authority from the central government to local authorities (Noah and Eckstein, 1992:147-170; Skolverket, 2005:6). This decentralisation process is characterised by the introduction of the curriculum called Lpo 94 in the 1990s, which emphasised a large local freedom to design education (Noah and Eckstein, 1993; SOU 2008:27). To summarise, I interpret England and Sweden as being towards the opposite ends of a scale measuring centralised and decentralised processes. With the two education systems being at opposite ends of this scale, it is inevitable that unintended and negative washback is recognised. It is conceivable that this will in turn deliver the motivation for the next round of reforms.

The current Swedish reform propositions made in SOU 2008:27, while featuring some centralising aspects such as the proposal to increase the number of nationally tested subjects, also aim to maintain teachers’ autonomy by allowing flexible delivery of such nationally designed tests. Whether these reforms will have an overall centralising effect or not, is therefore impossible to predict without seeing the implementation. In England, however, current reforms represent a further move towards centralised processes since they emphasise and propose a more unified qualifications framework (Green Paper, 2008; Department for children, schools and families, 2008).
The second aim of this thesis concerns the match between testing objectives and teachers’ attitudes. The objectives for the Swedish national testing system can be broadly summarised to comprise of: Grading assistance, Diagnostic Function, and National Monitoring. As it was difficult to find officially stated objectives for the English system of external examinations at GCSE and A level, I assumed that the English system may be measured against the same objectives. In general, respondents in both countries seemed to agree that a national testing system should meet such objectives but significant differences arose in whether they thought that the current system did actually meet them.

Firstly, English respondents, in contrast to their Swedish counterparts, do not think that the examination system meets the diagnostic objective. This is likely to stem from the inability in the English system to access useful formative information for individual pupils. Due to the local marking in the Swedish system, this information is readily available. The price, however, is that the marking has to be done locally.

Secondly, a similar picture arises with reference to the national monitoring objective. At first sight, the external and centralised marking approach taken in England seems to facilitate the implementation of consistent marking across the cohort and across years. Despite this, English respondents, in contrast to their Swedish counterparts, were divided on whether the national monitoring objective was met or not. The questionnaire highlighted unhappiness of English respondents with the external marking system which may undermine respondents’ perception of the national monitoring function. Furthermore, the division between English respondents could be due to the current public debate on how increasing average scores on examinations may arise from pupils choosing “easier” subjects (Smithers, 2006; Paton, 2007a). Perhaps surprisingly, Swedish teachers, despite the shortcomings of internally marked papers, believe that this function is attained on the whole.

With regards to the third aim, concerning teachers’ attitudes towards effects of testing, a number of interesting issues arose from the responses to the questionnaires. In particular, the English system (i) produces a mismatch of what teachers would like to see in the syllabus and what is tested and (ii) encourages a teaching delivery with undue focus on the tests themselves.

Respondents’ expressed preference for external marking is higher in England compared to Sweden. It was interesting to note that the Swedish respondents displayed a preference for external marking although the current system with internal marking is described in research (Skolverket, 2005) to be perceived rather well. It seems likely that teachers mainly think of the marking workload they could potentially offload, although the questionnaire did not highlight that particular issue. It is further interesting to pose the question as to whether this could be done without paying a price, in terms of achieving the current objectives for national tests, in particular the diagnostic function. My conclusion is that without a system of feeding back results of individual pupils, this objective is in jeopardy. In England where individual pupils’ results are not being forwarded to teachers, the perception of the respondents is that examinations currently do not serve any diagnostic function. Despite this, there is still a strong preference for external marking. It seems understandable to me that teachers, when considering a move to internal marking, would fear the extra workload.

From the questionnaire responses concerning effects of testing, I further draw the conclusion that the English respondents are discontent with negative effects generated by league tables. As a response to similar opinions by groups of teachers in England, the GTC has proposed a system where only a limited number of pupils’ results and schools are used for comparison. This is currently the case in Sweden for national tests on an upper secondary level (Skolverket, 2008b).
Judging from research done by Skolverket (2005:26) which indicated that little attention is paid to league tables in Sweden, I conclude that such a system could be beneficial in an English educational context. This would potentially address the concerns raised by English respondents with respect to the negative effects of league tables.

To summarise, the choice of words in the objectives for national tests in Sweden clearly illustrates how the Swedish testing system is fundamentally different from its English equivalent. It seems to me as if national tests in Sweden are supposed to support teachers in the grading process, whereas in England, the grading process is largely controlled by examinations, and teachers are mainly acting in a supporting role. It is not the purpose of this thesis to rate pros and cons or to express a preference for decentralisation or centralisation, nor to opt for a system of national tests or examinations. It is rather intended to be a reminder that countries with very different education and assessment culture can indeed learn from each other.

5.1 Future research

It would be interesting to see further research into areas where I failed to produce any conclusive results. One aspect could be to perform interviews with teachers, in addition to questionnaires in order to obtain a better insight into teachers’ views on summative testing. Moreover, it could be interesting to investigate teachers’ attitudes in a more specific context, e.g. attitudes towards specific subjects where national tests and examinations occur. Another aspect that could be interesting to investigate is if there are differences in attitudes between teachers of different subjects: science, language etc.

It could be interesting to investigate teachers’ reactions to and opinions of reforms currently being implemented in both countries. Finally a closer look at governing bodies’ motivation to steer education in certain ways and explore in what way these reforms relate to opinions and attitudes of teachers could be rewarding.
6 References

6.1 Published references


### 6.2 Electronic references


Ward, Helen & Shaw, Michael (2002). “End of the line for teaching to the test.” Published in *Times Education Supplement* on June 28th.
Appendices

Appendix 1

Hi there,
My name is Jessica Persson, in the middle of writing my thesis for my teaching degree in English as a Foreign Language and Physical Education at Umeå University, Sweden.

I kindly ask you to answer this 5-10 minute Online survey on teachers’ attitudes to testing in England compared to Sweden.

Your answers will be completely anonymous and will be of great help for my survey. Do not hesitate to contact me if you have any comments or questions.
attitudestotesting@gmail.com

Thank you!

Best Regards,
Jessica Persson
### 3. Do you think that external examinations at GCSE and A level CURRENTLY assist in setting fair and reliable grades (grading assistance)?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>62.3%</td>
<td>33</td>
</tr>
<tr>
<td>Don't know</td>
<td>3.8%</td>
<td>2</td>
</tr>
</tbody>
</table>

**answered question** 53  
**skipped question** 2

### 4. Do you think that external examinations at GCSE and A level SHOULD assist in setting fair and reliable grades (grading assistance)?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>100.0%</td>
<td>53</td>
</tr>
<tr>
<td>No</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Don't know</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

**answered question** 53  
**skipped question** 2

### 5. Do you think that external examinations at GCSE and A level CURRENTLY show students’ strengths and weaknesses (diagnostic function)?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>43.4%</td>
<td>23</td>
</tr>
<tr>
<td>No</td>
<td>49.1%</td>
<td>26</td>
</tr>
<tr>
<td>Don't know</td>
<td>7.5%</td>
<td>4</td>
</tr>
</tbody>
</table>

**answered question** 53  
**skipped question** 2
6. Do you think that external examinations at GCSE and A level SHOULD show students’ strengths and weaknesses (diagnostic function)?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>86.8%</td>
<td>48</td>
</tr>
<tr>
<td>No</td>
<td>5.7%</td>
<td>3</td>
</tr>
<tr>
<td>Don’t know</td>
<td>7.5%</td>
<td>4</td>
</tr>
<tr>
<td>Answered question</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>Skipped question</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

7. Do you think that external examinations at GCSE and A level CURRENTLY indicate the extent of students’ achievements in England (national monitoring)?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>45.3%</td>
<td>24</td>
</tr>
<tr>
<td>No</td>
<td>47.2%</td>
<td>25</td>
</tr>
<tr>
<td>Don’t know</td>
<td>7.5%</td>
<td>4</td>
</tr>
<tr>
<td>Answered question</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>Skipped question</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

8. Do you think that external examinations at GCSE and A level SHOULD indicate the extent of students’ achievements in England (national monitoring)?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>90.6%</td>
<td>48</td>
</tr>
<tr>
<td>No</td>
<td>5.7%</td>
<td>3</td>
</tr>
<tr>
<td>Don’t know</td>
<td>3.6%</td>
<td>2</td>
</tr>
<tr>
<td>Answered question</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>Skipped question</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>
9. Would you prefer alternative methods for assessing students’ achievements instead of external examinations?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>27.1%</td>
<td>13</td>
</tr>
<tr>
<td>No</td>
<td>62.5%</td>
<td>30</td>
</tr>
<tr>
<td>Don’t Know</td>
<td>10.4%</td>
<td>5</td>
</tr>
</tbody>
</table>

If yes, what kind of alternative method(s) 

Answered question: 48

Skipped question: 7

10. Currently, English students are subjected to external examinations at GCSE and A level. To what extent do you agree or disagree with each of the following? External examinations .......

<table>
<thead>
<tr>
<th></th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don’t know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>.... should begin at GCSE level</td>
<td>45.7% (16)</td>
<td>37.1% (13)</td>
<td>8.6% (3)</td>
<td>8.6% (3)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>35</td>
</tr>
<tr>
<td>.... should begin at A level</td>
<td>7.0% (3)</td>
<td>16.3% (7)</td>
<td>14.0% (6)</td>
<td>41.8% (18)</td>
<td>10.6% (6)</td>
<td>2.3% (1)</td>
<td>43</td>
</tr>
<tr>
<td>.... should not happen at all</td>
<td>0.0% (0)</td>
<td>4.2% (1)</td>
<td>12.5% (3)</td>
<td>33.3% (8)</td>
<td>56.0% (12)</td>
<td>0.0% (0)</td>
<td>24</td>
</tr>
<tr>
<td>.... should only be taken in a core of compulsory subjects at A level</td>
<td>0.0% (0)</td>
<td>16.7% (7)</td>
<td>14.3% (6)</td>
<td>31.0% (13)</td>
<td>36.7% (15)</td>
<td>2.4% (1)</td>
<td>42</td>
</tr>
</tbody>
</table>

Answered question: 48

Skipped question: 7
11. Currently examinations in England are marked by externally appointed markers, which is referred to as external marking in the question below. In Sweden, national tests are marked by teachers in schools, which is referred to as internal marking. To what extent do you agree or disagree with each of the following statements?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don't know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>External marking is consistent across the student cohort/objective</td>
<td>14.7% (5)</td>
<td>26.5% (9)</td>
<td>11.8% (4)</td>
<td>36.3% (12)</td>
<td>11.8% (4)</td>
<td>0.0% (0)</td>
<td>34</td>
</tr>
<tr>
<td>The administration of external marking is inefficient</td>
<td>2.2% (1)</td>
<td>42.2% (19)</td>
<td>20.0% (9)</td>
<td>24.4% (11)</td>
<td>11.1% (5)</td>
<td>0.0% (0)</td>
<td>45</td>
</tr>
<tr>
<td>Internal marking would be consistent across the student cohort</td>
<td>9.1% (3)</td>
<td>10.2% (8)</td>
<td>9.1% (3)</td>
<td>45.6% (15)</td>
<td>18.2% (6)</td>
<td>0.0% (0)</td>
<td>33</td>
</tr>
<tr>
<td>Internal marking would be affected by the marker's knowledge of the students' identity</td>
<td>33.3% (15)</td>
<td>50.0% (24)</td>
<td>8.3% (4)</td>
<td>6.3% (4)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>46</td>
</tr>
</tbody>
</table>

answered question 48

skipped question 7

12. Overall, for examinations at GCSE and A level, would you prefer internal or external marking?

<table>
<thead>
<tr>
<th>Marking Type</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal marking</td>
<td>12.5%</td>
<td>6</td>
</tr>
<tr>
<td>External marking</td>
<td>87.5%</td>
<td>42</td>
</tr>
</tbody>
</table>

answered question 48

skipped question 7
13. With regard to the effects of testing, to what extent do you agree or disagree with the following statements?

<table>
<thead>
<tr>
<th>Statement</th>
<th>agree</th>
<th>disagree</th>
<th>know</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination results give an accurate evaluation of student learning</td>
<td>64.5% (20)</td>
<td>12.9% (4)</td>
<td>26.7% (12)</td>
<td>31</td>
</tr>
<tr>
<td>Examination helps me evaluate my teaching</td>
<td>44.4% (20)</td>
<td>13.3% (6)</td>
<td>26.7% (12)</td>
<td>45</td>
</tr>
<tr>
<td>Examination helps me identify areas of improvement for my students</td>
<td>50.0% (15)</td>
<td>13.3% (4)</td>
<td>26.7% (8)</td>
<td>30</td>
</tr>
<tr>
<td>Examination helps my assessment of students</td>
<td>64.4% (29)</td>
<td>13.3% (6)</td>
<td>6.7% (3)</td>
<td>45</td>
</tr>
<tr>
<td>Examination creates a lot of pressure on teachers</td>
<td>46.7% (21)</td>
<td>2.2% (1)</td>
<td>8.9% (4)</td>
<td>45</td>
</tr>
<tr>
<td>Examination creates a lot of pressure on students</td>
<td>33.3% (10)</td>
<td>26.7% (8)</td>
<td>6.7% (2)</td>
<td>30</td>
</tr>
<tr>
<td>I feel there is a discrepancy between what I think should be taught and what the examination tests emphasizes</td>
<td>33.3% (10)</td>
<td>26.7% (8)</td>
<td>6.7% (2)</td>
<td>30</td>
</tr>
<tr>
<td>The content of my teaching is largely determined by the examinations</td>
<td>51.1% (23)</td>
<td>37.8% (17)</td>
<td>8.9% (4)</td>
<td>45</td>
</tr>
<tr>
<td>Examinations should be used for comparing schools</td>
<td>43.3% (13)</td>
<td>26.7% (8)</td>
<td>0.0% (0)</td>
<td>30</td>
</tr>
<tr>
<td>I know how well my school performs on examinations</td>
<td>45.5% (20)</td>
<td>9.1% (4)</td>
<td>4.5% (2)</td>
<td>44</td>
</tr>
<tr>
<td>My school seems more interested in high test scores than improving overall student learning</td>
<td>40.0% (12)</td>
<td>26.7% (8)</td>
<td>13.3% (4)</td>
<td>30</td>
</tr>
</tbody>
</table>

**answered question** 45

**skipped question** 10

---

Page: WANT A SUMMARY OF MY FINDINGS?

14. If you would like to read the summary of my findings, then write your e-mail in the space below.

<table>
<thead>
<tr>
<th>Response</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>view</td>
<td>12</td>
</tr>
</tbody>
</table>

**answered question** 12

**skipped question** 43
Appendix 2

Hej,
Jag heter Jessica Persson, och håller på att skriva mitt examensarbete på lärarutbildningen vid Umeå Universitet.

Klicka helt enkelt på denna länk för att komma till enkäten:
http://www.surveymonkey.com/s.aspx?sm=H_2f3k7qc2aqsXJB_2fmhtvYAg_3d_3d

Dina svar är helt anonyma och kommer vara till stor hjälp för min undersökning.
Vidarebefordra gärna denna länk till andra lärare och tveka inte att kontakta mig om du har kommentarer eller frågor.
attitudestotesting@gmail.com

Tack!
Med vänlig hälsning,
Jessica Persson

<table>
<thead>
<tr>
<th>Page: INTRODUCTION</th>
</tr>
</thead>
</table>

1. What age group do you primarily teach?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>primary (tidigare delen av grundskolan)</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>secondary (senare delen av grundskolan)</td>
<td>25.0%</td>
<td>4</td>
</tr>
<tr>
<td>upper secondary (gymnasiet)</td>
<td>75.0%</td>
<td>12</td>
</tr>
<tr>
<td>other</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

If other (please specify) 0
answered question 16
skipped question 0

2. What subject(s) do you primarily teach?

<table>
<thead>
<tr>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
</tr>
</tbody>
</table>

answered question 16
skipped question 0
### Page: OBJECTIVES OF EDUCATIONAL AUTHORITY

3. Below is a list with the objectives for the Swedish national tests (Nationella prov). Do you think that the national tests achieve the following objectives? Syftet med det nationella provssystemet skall vara att:

<table>
<thead>
<tr>
<th>Objective</th>
<th>Yes</th>
<th>No</th>
<th>Don’t Know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>- bidra till okad måluppfyllelse för eleverna</td>
<td>40.0% (6)</td>
<td>53.3% (8)</td>
<td>6.7% (1)</td>
<td>15</td>
</tr>
<tr>
<td>- visa på elevers starka och svaga sidor (diagnostic function)</td>
<td>86.7% (13)</td>
<td>6.7% (1)</td>
<td>6.7% (1)</td>
<td>15</td>
</tr>
<tr>
<td>- konkretisera kursmål och betygsriter</td>
<td>60.0% (9)</td>
<td>33.3% (5)</td>
<td>6.7% (1)</td>
<td>15</td>
</tr>
<tr>
<td>- stödja en likvärdig och rättvis bedömning och betygsättning (grading assistance)</td>
<td>73.3% (11)</td>
<td>20.0% (3)</td>
<td>6.7% (1)</td>
<td>15</td>
</tr>
<tr>
<td>- ge underlag för en analyser i vilken utsträckning kunskapsmålen nås (national monitoring)</td>
<td>86.7% (13)</td>
<td>0.0% (0)</td>
<td>13.3% (2)</td>
<td>15</td>
</tr>
</tbody>
</table>

**Answered question** 15

**Skipped question** 1

### Page: TYPE OF ASSESSMENT/ MARKING

4. Would you prefer alternative methods for assessing students’ achievements instead of national tests?

<table>
<thead>
<tr>
<th>Response</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>14.3%</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>57.1%</td>
<td>8</td>
</tr>
<tr>
<td>Don’t know</td>
<td>28.6%</td>
<td>4</td>
</tr>
</tbody>
</table>

If yes, what kind of alternative methods?

**Answered question** 14

**Skipped question** 2
5. Currently, Swedish students are subjected to national tests in year 5, year 9 and in upper secondary school. To what extent do you agree or disagree with each of the following statements? National testing should ...

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don’t know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>- begin at an earlier stage than year 5</td>
<td>7.1% (1)</td>
<td>7.1%  (1)</td>
<td>14.3%   (2)</td>
<td>35.7%    (5)</td>
<td>28.5%  (4)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>- begin at year 5</td>
<td>0.0% (0)</td>
<td>21.4% (3)</td>
<td>71.4% (2)</td>
<td>14.3%    (2)</td>
<td>7.1%  (1)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>- begin at year 9</td>
<td>14.3% (2)</td>
<td>21.4% (3)</td>
<td>7.1% (1)</td>
<td>21.4%    (3)</td>
<td>28.6%  (4)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>- begin in upper secondary school (gymnasiet)</td>
<td>14.3% (2)</td>
<td>7.1% (1)</td>
<td>21.4% (3)</td>
<td>14.3%    (2)</td>
<td>35.7%  (5)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>- not happen at all</td>
<td>7.1% (1)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>28.6%    (4)</td>
<td>50.0%  (7)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>- be confined to Swedish, Swedish as a second language, English, and</td>
<td>15.4% (2)</td>
<td>15.4% (2)</td>
<td>38.5% (5)</td>
<td>7.7%    (1)</td>
<td>7.7%  (1)</td>
<td>15.4% (2)</td>
<td>13</td>
</tr>
<tr>
<td>Mathematics.</td>
<td>7.1% (1)</td>
<td>28.6% (4)</td>
<td>28.6% (4)</td>
<td>7.1%    (1)</td>
<td>21.4%  (3)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
</tbody>
</table>

answered question 14

skipped question 2

6. Currently national tests in Sweden are marked by teachers in schools, which is referred to as internal marking in the question below. In many other countries national tests and exams are marked by externally appointed markers (people outside the school), which is referred to as external marking. To what extent do you agree or disagree with each of the following statements?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don’t know</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Internal marking of national tests is consistent across the student cohort</td>
<td>7.1% (1)</td>
<td>14.3% (2)</td>
<td>21.4% (3)</td>
<td>21.4% (3)</td>
<td>0.0% (0)</td>
<td>35.7% (5)</td>
<td>14</td>
</tr>
<tr>
<td>- Internal marking of national tests is affected by the marker’s knowledge of the students’ identity</td>
<td>0.0% (0)</td>
<td>42.9% (6)</td>
<td>28.6% (4)</td>
<td>14.3% (2)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>- External marking of national tests is consistent across all pupils</td>
<td>14.3% (2)</td>
<td>21.4% (3)</td>
<td>28.6% (4)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>35.7% (5)</td>
<td>14</td>
</tr>
</tbody>
</table>

answered question 14

skipped question 2

7. Overall, for the national test, would you prefer internal or external marking?

<table>
<thead>
<tr>
<th>Preference</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>External marking</td>
<td>57.1%</td>
<td>8</td>
</tr>
<tr>
<td>Internal marking</td>
<td>42.9%</td>
<td>6</td>
</tr>
</tbody>
</table>

answered question 14

skipped question 2
8. To what extent do you agree or disagree with the following statements?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Don’t know</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>National test results give an accurate evaluation of student learning</td>
<td>0.0% (0)</td>
<td>42.9% (6)</td>
<td>42.9% (6)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>National tests help me evaluate my teaching</td>
<td>14.3% (2)</td>
<td>50.0% (7)</td>
<td>14.3% (2)</td>
<td>7.1% (1)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>National tests help me identify areas of improvement for my students</td>
<td>14.3% (2)</td>
<td>71.4% (10)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>National tests help my assessment of students</td>
<td>7.1% (1)</td>
<td>78.6% (11)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>National testing creates a lot of pressure on teachers</td>
<td>14.3% (2)</td>
<td>42.9% (6)</td>
<td>21.4% (3)</td>
<td>7.1% (1)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>National testing creates a lot of pressure on students</td>
<td>21.4% (3)</td>
<td>57.1% (8)</td>
<td>7.1% (1)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>I feel there is a discrepancy between what I think should be taught and what the tests emphasises</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>35.7% (5)</td>
<td>42.9% (6)</td>
<td>0.0% (0)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>Marking national tests take up too much of my time</td>
<td>21.4% (3)</td>
<td>28.6% (4)</td>
<td>28.6% (4)</td>
<td>14.3% (2)</td>
<td>0.0% (0)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>The content of my teaching is largely determined by the national tests</td>
<td>0.0% (0)</td>
<td>7.1% (1)</td>
<td>21.4% (3)</td>
<td>50.0% (7)</td>
<td>7.1% (1)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>National tests should be used for comparing schools</td>
<td>0.0% (0)</td>
<td>28.6% (4)</td>
<td>35.7% (5)</td>
<td>14.3% (2)</td>
<td>14.3% (2)</td>
<td>7.1% (1)</td>
<td>14</td>
</tr>
<tr>
<td>I know how well my school performs on national tests</td>
<td>7.1% (1)</td>
<td>42.9% (6)</td>
<td>14.3% (2)</td>
<td>21.4% (3)</td>
<td>0.0% (0)</td>
<td>14.3% (2)</td>
<td>14</td>
</tr>
<tr>
<td>My school seems more interested in high test scores than improving overall student learning</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>7.1% (1)</td>
<td>42.9% (6)</td>
<td>28.6% (4)</td>
<td>21.4% (3)</td>
<td>14</td>
</tr>
</tbody>
</table>

9. If you would like a summary of my findings, then please write your e-mail address; otherwise, leave it blank

<table>
<thead>
<tr>
<th>Email Address</th>
<th>Response Count</th>
</tr>
</thead>
</table>