Degree Project

Up to standard?
A CEFR-related comparative study of Swedish and Norwegian model texts for assessing the national exam in written English for 9th graders

Author: Simon Almqvist
Supervisor: Charlotte Hommerberg
Examiner: Ibolya Maricic
Date: 24 August 2019
Subject: English
Level: Undergraduate
Course code: 2UV90E
Abstract

This study aims at exploring the quality of the Swedish and Norwegian national tests using their respective model texts for assessing. The study does so by relating them to the CEFR and the grading tool Write & Improve within the context of the two countries and the field of language testing. The study finds there to be a set of inconsistencies between what the national tests want to do and what they actually do. In particular, the study finds the Swedish national test not to be up to its own standards.

Keywords

National test, Sweden, Norway, model texts, language testing, Write & Improve, CEFR

Thanks

My sincerest thanks to Charlotte Hommerberg, Ibolya Maricic, Elena de Wachter and Sofia Gomes for their invaluable help and feedback.
Contents

1 Introduction ......................................................................................................................... 1

2 Context and Theory ........................................................................................................... 4
   2.1 National tests in Sweden ............................................................................................... 4
   2.2 National tests in Norway ............................................................................................. 8
   2.3 Language testing .......................................................................................................... 10

3 Method and Material ......................................................................................................... 15
   3.1 Material ....................................................................................................................... 15
   3.2 Method ......................................................................................................................... 22

4 Results ................................................................................................................................ 24
   4.1 Write & Improve .......................................................................................................... 24
   4.2 Further examining of N2 and SE- ................................................................................ 26
   4.3 EVP results .................................................................................................................. 29

5 Analysis and Discussion .................................................................................................... 30
   5.1 N2 and SE- .................................................................................................................. 30
   5.2 Comparison of results ................................................................................................. 32
   5.3 Validity of Write & Improve and Other Considerations ............................................. 35
   5.4 Contextual consequences ............................................................................................ 39

6 Conclusion ........................................................................................................................ 43

Reference list ......................................................................................................................... 46

Appendices ............................................................................................................................ I
   Appendix A - Model texts ................................................................................................. I
   Appendix B – Commentary on model texts from within the document of guidelines XI
1 Introduction

In May of 2017 I was interning at an upper elementary school in Sweden, where I corrected national tests along with a group of experienced English teachers at the school. I had been handed guidelines to read a few days before, as a means of helping me assess the pupils’ texts with a national standard. We were sitting at a big table working through the texts and talking back and forth when I interrupted the conversation and questioned the quality of the model texts (previous graded answers from students with the same assignment) we had read in our guidelines, in particular the model text that was meant to represent the grade E (lowest pass). It did not match my initial expectations nor my interpretation of the criteria for a pass at that level (9th grade). All the teachers gave the same dejected look of resignation - neither did it match the standard they had set for their pupils during the course of three years.

That experience was perhaps more of a shock considering the context – it was Sweden. Swedes pride themselves on being one of the best non-native English speakers in the world. In fact they are currently ranked in first place in the English Proficiency Index, at the top of the chart along with all the other Nordic countries (www.ef.se/epi/). For this reason, our expectations on the texts that were meant to represent passing grades were high, which led to disappointment. This contradiction invites speculation. If the number one on EF’s list (Sweden) indeed has a low set bar for national exams, how does that figure? Is it a fluke or can the same be said for other countries at the top? To answer these questions one can glance at the top and notice all the Nordic countries are placed high up. After some further research Norway proved to be a good candidate to compare with as Denmark was ruled out due to the inaccessibility of comparable material. The aim of this study is to try to understand the seemingly low quality of the model texts in the Swedish national exam in English and by extension also find out if the criticism holds up. The study does not, before-hand, entertain any idea or suspicion as to why or how the texts appear poor. For that reason, this study will proceed with an open-minded approach, meaning that the subject will be tackled from different angles: within the context of language testing, with references from both test producers and critics, in comparison to a similar, neighboring country, and finally, through a detailed analysis of the language in the texts. This will be reviewed through the lens of the Common European Framework of Reference for Language and with the help of an evaluative software that is based on said framework.
These are this study’s research questions:

*How do the Swedish model texts in written English compare to the Norwegian ones in terms of context (background and structure) and language (formulations, assessment and CEFR)?*

*What kind of results are produced by a software like Write & Improve and can those results be considered accurate – i.e. does it have validity?*

The Common European Framework of Reference for Language (henceforth CEFR) is, as its name suggests, a linguistic framework within Europe that has been carefully developed by the 41 member states of the Council of Europe for educational purposes. Among other things, it has certain proficiency levels meant to be transferable throughout the continent (Council of Europe 2001). An Italian’s proficiency in a language would thus be comparable to that of a Belgian’s. These levels of proficiency are as follows, from least proficient to most: A1, A2, B1, B2, C1, C2. As can be seen in the “Structured overview of all CERF scales” (Council of Europe 5) basic users are A level and can understand simple and common concepts and communicate on familiar topics. Independent users are B level and can understand the main idea of complex texts, can interact with a degree of fluency and produce clear, detailed text on various topics. Proficient users are C level, meaning that they can express themselves with great precision and variation and have little or no difficulty understanding anything (Council of Europe 5-7). The above descriptions are based on the can-do statements the CEFR provides and do not completely represent the 6 levels, as there is a difference in proficiency within the respective levels too (i.e. A1 is not the same as A2).

The CEFR also supplies a standardized terminology to better understand different language levels and more easily describe them (Cambridge UP 2013:9), rather than one country saying one thing and the other country another.

A comparative study like this would, should one pursue it holistically (that is to look at the text as a whole and decide the general qualities of it), require many examiners’ point of view as this study’s attempt at a holistic understanding of the texts would not suffice.
The experience that gave birth to this study was the seemingly low quality of the model texts for the national tests in Sweden. When two English teachers in Sweden (skolvarlden.se) also criticized the Swedish National Agency of Education (SNAE) (In Swedish, Skolverket) for setting the bar too low on national tests, they reaffirmed and brought more attention to the issue. More specifically, the written English national test was brought into question. Bruun & Diaz claim in the aforementioned article that SNAE want to adhere to the CEFR-levels of proficiency, but that the aforementioned model texts that are given to teachers to help them discern the levels of proficiency required for each grade, do not hold up to these required standards. The required standard being that 9th graders at least show a B1 proficiency in English. Bruun and Diaz used a software called Write & Improve to help determine the CEFR level of proficiency and based their criticism on those results. This naturally raised the second research question about the validity of the software.

For this reason, the CEFR proficiency levels are being applied in the analysis of the Swedish and Norwegian model texts which makes the model texts more comparable, not only to each other, but also on an international level, which is academically significant in the sense that it broadly invites further research. On that note of comparability, Sweden’s western cousins in Norway have a similar system where the examiners are internal rather than external, which is a common international practice (Sundqvist et al. 2017:1). Moreover, the two countries share cultural and historical bonds and are socioeconomically comparable. Even more interesting are the slight differences that exist on the subject of national testing, and the potentially meaningful conclusions that could thus be drawn from a comparative study of the model texts. An example of these slight differences might include the fact that Norway uses digital tests. They also use in-house teachers to assess the tests, but they are in that role called “sensors”, receive a special education and can actively take part in the creation of the tests, according to Matilda Burman, test assessment manager at the The Norwegian Directorate for Education and Training (Personal communication). Nonetheless, when actually sitting down to assess, model texts are used as a key point of assessment example (Personal communication with Matilda Burman). As such, the language shown and the decisions that led to those particular model texts being used are influential to the assessment process and for that reason important to observe. More vital differences will be portrayed in sections 2.1, 2.2 and 3 of this essay.
While the world is a big context, the Nordic region is a more defined one. The Nordic countries all display high levels of English proficiency and have been working to differing degree toward integrating the CEFR into policy. Despite these efforts, Erickson & Pakula (2017) have found, in their study “The common european framework for language: Learning, teaching, assessment – a nordic perspective” (2017: 18-19) (my translation) that the degree to which policy has been put into practice by teachers is questionable at best. They also state that Norway has, just like Sweden, been working towards making their grades relatable to CEFR-levels but not to the same degree (2017: 13-17).

Consideration: Since Write & Improve will be an important tool used to cover language criteria of the first research question, the questions become inevitably intertwined. This means it potentially sets itself up to either success in both regards or failure in both regards. Should Write & Improve prove a valid tool for language testing, the comparison will lend itself to that same validity – generating reliability. On the other hand, if it turns out to be too flawed, the comparison of the model texts will also suffer the consequences. However, a failure in regard to Write & Improve would not be a failure for this paper, because then it can at least conclude that the software is not sufficient for language testing, when one peeks behind its helpful façade. The software is not originally meant to be used in this way, but rather to help students of English improve, but as it has been used to criticize language quality by Bruun & Diaz (2017), the software is thrown into a looming debate where curiosity leads to scrutiny.

2 Context and Theory

According to Erickson and Pakula (2017), the Nordic countries’ devotion to CEFR often is not applied in practice and the extension to which it is applied is largely based on individual schools and/or teachers own initiative. Ergo, when we use the CEFR as reference to tackle the issue of text quality in the national tests, we could potentially also gain an insight into how well CEFR has been adopted – to what extent principle has become reality (Erickson & Pakula 2017:15, 18-9)

2.1 National tests in Sweden

For most of the 20th century the Swedish school system was governed centrally, without much opposition. However, in the 1970s criticism arose that questioned the extent of
regulation and bureaucracy. This criticism led to change and by the end of the 1980s the Swedish school, although still state-regulated, had been decentralized, deregulated and largely so with the influence of New Public Management (NPM). NPM is a term often used for the idea that public departments and institutions can and should be run in the same practical manner that private companies are run (Arensmeier & Lennqvist 2017: 49-53).

Another development that ran parallel to this, however delayed, was the transition from a trust-based policy toward teachers to a mistrust-based policy. This is most clearly emphasized by the increased quantity and value of standardized testing and quality controls and according to Arensmeier & Lennqvist (2017: 49-53) this development only really took off in the 2000s. This increase in quality controls and monitoring Montin claims is widespread in society and applies to more institutions than the educational one (Montin 2015: 59-61). Moreover, Arensmeier & Lennqvist find that the increased controls reinforce the idea that teaching is a semi-profession (Arensmeier & Lennqvist 2017: 49).

One such attempt of controlling and measuring knowledge is through standardized tests and the official definition of the national tests in Sweden and their purpose is expressed on the website of SNAE:

The purpose of the national tests is mainly to support an equal and fair assessment and grading, and supply the basis of analysis regarding the degree to which knowledge goals are being accomplished on a school level, a ministry level and on a national level. The national tests are not final examinations rather a part of the teachers’ gathered information of a pupil’s knowledge. It is the government that decides which subjects, years and contexts that national tests will be held. (SNAE, own translation)

This description implies that without the national tests foreign language teachers are unable to grade reliably and confirms the control mechanism discussed earlier in monitoring how goals are being met. The Swedish Schools Inspectorate (In Swedish, Skolinspektionen) have reported that the standardized national tests are lacking in reliability because of the teachers who assess it. The teachers working at the school more often than is desired award a different grade than other teachers re-marking the tests would. (Skolinspektionen 2013). Consequently, Swedish teachers’ credibility when
it comes to correctly assessing the national tests has been somewhat tarnished as of recent (Gustafsson & Erickson 2013, Sundqvist et. al 2017). This issue is highly relevant to contemporary Swedish pedagogics.

Sundqvist et al. (2017:6), in turn, observe that Foreign language teachers have a unique understanding of the effect and implementation of the tests, yet do not directly take part in the formulation of the tests – something she finds to be problematic. Furthermore, the national test is accompanied by “detailed guidelines” (SNAE evaluation) for teachers to follow. This fits well into the idea of mistrust-driven policies, since if the teachers were trusted to assess fairly such extensive tools for coherency would not be needed. This mistrust can sometimes become double-edged however, as shown by the aforementioned teacher-published article by Bruun & Diaz in Skolvärlden (2017). In the detailed guidelines supplied by SNAE there are model texts featured meant to perfectly correspond to student works of different levels. There is one text for F (fail), one for E-, one for E+, one for D- and so on. Bruun & Diaz applied these model texts to the program Write & Improve to determine their level of CEFR proficiency for English using the six levels of CEFR proficiency. SNAE claims to have based their knowledge goals for English and Modern languages upon this very same European framework. (SNAE 2012) Despite this, Bruun and Diaz’ test with Write & Improve found that the three texts they used, which are meant to be the equivalent of a pass in 9th grade (E, or in CEFR, B1), at highest reach the level of A2. This would be the expected level of proficiency in grades 6 and 7 in CEFR, according to SNAE (see Figure 1 on page 4)(Skolvärlden 2017). The University producing the national tests in Sweden has, along with European experts, stated that the national tests in fact do correspond to the CEFR grades, going against the initial observations of this paper. (Broek & van den Ende 2013: 48)

The apparent clash between the official statement above and the initial findings of this paper serves to highlight how important it is that this matter is scrutinized. Upon further investigation and a phone call with Sara Bruun, one of the authors of the article, a near-consensus can be deciphered as she describes her broad experience of discussing with fellow teachers and claiming that she is “yet to find an English teacher who finds that the model texts are up to standard”(Personal communication, Sara Bruun). This does not necessarily mean that Swedish teachers generally agree with her as no detailed study into that has been done. It does, however, serve as a strong indication and is something that this paper will examine (Personal communication with Sara Bruun).
The work with CEFR in Sweden goes far back as the country participated actively in the 1970s to 80s on the European Council’s proceedings on language within Europe. As a consequence, the European perspective was largely present domestically in the 1980s and 90s as three curriculum reforms were implemented in 1980, 1982 and 1994 respectively. The CEFR began to be used for the development of national tests in English in 1998 and efforts were made to educate Swedish teachers on the matter, through national teacher conferences and published commentary material (Erickson & Pakula 2017:15-6). According to Erickson & Pakula the translation of CEFR into Swedish in 2009 was done parallel to further informational campaigning, culminating in SNAE “explicitly” connecting the steps for the subject of English to the CEFR 6-grade scale in the most recent curriculum reform, that of 2011 (2017:16).

As shown in Figure 1 featured below, the expected levels of proficiency in CEFR (GERS in the Swedish translation) correlate with different years in the Swedish school. Most notably, for this study, year 9 (Åk 9) has the expected proficiency level B1.

<table>
<thead>
<tr>
<th>GERS, nivå</th>
<th>A 1.1</th>
<th>A 1.2</th>
<th>A 2.1</th>
<th>A 2.2</th>
<th>B 1.1</th>
<th>B 1.2</th>
<th>B 2.1</th>
<th>B 2.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEG</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Engelska</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grundskolan</td>
<td>åk 6</td>
<td>åk 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gymnasieskolan, kurs</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderna språk &amp; teckenspråk för hörande</td>
<td>Elevens val åk 9</td>
<td>Språkval åk 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grundskolan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gymnasieskolan, kurs</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. The correlation between CEFR-levels (here ”GERS”) and the progression in Swedish schools.*

Figure 1 is most notably showing that for year 9 (Åk 9”) the expected level is B1. Commentary material to the curriculum for English (Original title: Kommentarmaterial till kursplanen i engelska (Skolverket 2011:7). However, despite substantial governmental efforts the actual implementation of CEFR remains questionable with varying accounts of usage and a lack of systematic statistics (Eriksson 2017:17). Nevertheless, it has been made clear that the national test in written English and its assessment guidelines are meant to be corresponding to the European proficiency levels.
2.2 National tests in Norway

Unlike the Swedish system, the Norwegian one separates national tests from the regular education. They are linked in the sense that they have the same curriculum and teachers are meant to help students preparing for them. But the national tests in Norway are assessed by specially educated sensors. The principals of schools recommend teachers suitable for the position of sensor to the regional administrative school. According to the Norwegian Directorate for Education and Training (NDET) this improves the teachers’ ability to assess even after being a sensor (Interview, Matilda Burman). Additionally, they result in a separate grade from the one the teachers set (udir.no) and the quality of the tests is perceived as high according to the Norwegian sensors themselves, the most recent survey showing only 3.3 % found the test to fail in giving the opportunity for pupils to show different competences (NDET 2017:5).

On the Norwegian perspective on foreign language and language learning, the CEFR was influential even before its official release in 2001 (Erickson & Pakula 2017:13). The CEFR was translated to Norwegian in 2011 and has long been used as a point of reference for the national tests in English there. They further state that these national tests were developed at the University of Bergen and that they are “implicitly related” (2017:14 own translation) to the CEFR, both in terms of content and level (2017:14). Both Sweden and Norway have worked with the Common European Framework of Reference for languages: Learning, Teaching, Assessment (CEFR) although according to the NDET, Sweden has come further in assimilating to a European standard. Norway is, however, working on renewing their image on this according to udir.no/fagfornyelsen/.

The image as presented by the NDET comes across as modest and underwhelming when considering the more positive one presented by Erickson & Pakula (2017). This could be explained by the comparison to the extensive Swedish embracement and speculatively the fact that, for the purposes of this study, they were contacted by a Swede. Nevertheless, it is fair to say Norwegian national tests in English and the CEFR proficiency levels are intertwined and thus relevant to this study.

This process of intertwinement began in 2006, when an educational reform was implemented in which CEFR was considered an essential document for the English and Modern languages. This meant that the curriculum was greatly influenced by the CEFR, however without being mentioned explicitly in either the formulation of aims or
description of proficiencies (Erickson & Pakula 2017:13-16). Although that is true, many references have been made to the European documents in Norway and the country has seen increased use of self-evaluation since the enactment of the 2006 reform (Erickson & Pakula 2017:15). Some critique to this ambivalence (of using CEFR but not mentioning it explicitly) has arisen and might be accommodated in an upcoming reform. Curriculums are expected to be implemented by 2019. According to Erickson & Pakula, such changes and correlating efforts could improve comparability and studies like the one this paper is attempting (2017:15). Such efforts would improve external insight and measurability of both countries’ adoption of CEFR.

Previous findings by Erickson & Pakula casts a shadow on that prediction, though. As mentioned in the introduction, the Nordic countries all show will and commitment to integrating CEFR but the success is very varied (2017:18-9) with individual schools, universities and teacher’s initiatives/lack of initiatives being the most notable factors. To put it simply, Erickson & Pakula argue that Sweden and Norway promote the use of a European comparable standard but do not effectively put it into practice (2017:13-19). Broek & van den Ende argue for the same thing in “The Implementation of the Common European Framework for Languages in European Education Systems” (2013) when they state that “in Sweden, the main problem concerning the implementation of the CEFR is that teachers do not really use the CEFR as a tool in the classroom”. They also conclude that this lack of use affect elementary schools not only directly but also indirectly as the same problem is found in university teacher training as well, which affects the students becoming teachers and by extension the education that they eventually carry out. (Broek & van den Ende, 2013:61-64).

Another discussion that was sparked by the Norwegian educational reform of 2006 and is still ongoing is that of washback-effects. The 2006 Norwegian educational reform brought with it a lot of changes, most notably the structure of the national exam was changed and the curriculum was completely overhauled from a direct, instructive approach to more loosely defined “learning aims” (Ellingsund 2009:1). While these two changes had the obvious direct effect of changing the shape of the education and national tests, Ellingsund claims they also had an indirect, consequential effect, a so called “washback”-effect (2009:1-3). An illustration of the washback-effect is how the priorities of the Norwegian guideline document (see more in Section 3) indirectly affect teachers’ methods. The assignments that are featured in that document are followed by a figure divided into three categories. These categories are the aspects highlighted and
meant to look at when assessing the national tests. One of them is content (Guidelines Nor 2014:2), which, perhaps in traditional EFL (English for foreign learners) is not regarded as important as mastering the language but more on that note later. The point of this example is that the teachers/sensors that use this document as they assess national tests will bring that format with them back to regular school education and indirectly be affected by the focus of the national tests – thus shaping the regular education accordingly (Ellingsund 2009:2). As such, the washback-effect highlights the importance of formatting and language in the guidelines since it can influence teaching in the classrooms. For this reason, if the standardized exams are not properly formulated it could have a larger negative impact. On the other hand, the same could be said for “good” formulations having a larger positive impact.

Norway is far from alone in the endeavour of integrating the CEFR and Erickson & Pakula found that CEFR has had a growing impact on assessment in language education in Europe and even outside of Europe (2017:1). Erickson & Pakula mainly find the reception of CEFR to be largely positive, highlighting its functional and comparable strengths (2017:7). However, despite this established authority of the framework there has been outspoken criticism against it for being too linguistically focused, for not taking culture and interculturality within individual nations into account. In response to this criticism the letter R in the abbreviation CEFR has been emphasized, meaning it is meant as a point of reference and how different national and cultural entities comply with it is to some degree up to them. Another aspect of CEFR that has sparked debate is issues of interpretation. This means that countries too easily can interpret it in their way and loosely apply it. Without empirical substantiation they could claim that their curriculum and syllabi are in perfect alignment with the proficiency levels (2017:6-7), which begs the question if that’s what Sweden is doing.

As a response to this criticism, the European Council supplied a manual to aid in correctly interpreting proficiency levels called the “Manual for Relating Language Examinations to the Common European Framework of Languages” (CoE 2009). This study will not delve further into such a manual, however, since the aim of it is not primarily to study the wheels and cogs of the machine, but rather the results.

2.3 Language testing
How one goes about examining representative, model texts differs greatly due to the nature of language and humanity. Humans are not machines and will not give the same answer to a question every time. Even when a test boasts of reliability (roughly meaning it upholds a desired consistency) (Shillingsburg 2016:5-6), we can’t be assured that the teacher, sensor or examiner of that test has the same reliability in assessing text by text, case by case. A poignant example are parole judges, who are meant to always follow and apply the same principles of law. Yet a parole hearing after the judge has had lunch is 2 to 6 times more likely to result in a release than before lunch (Danziger et. al 2011). Testing language is of course not the same as a parole hearing but it can be argued that the same psychological effects can be at play, extenuating factors.

On the other hand, computerized language testing like the one being done in Write & Improve, has an issue of validity. How can a pre-programmed software grasp the width of a language with, according to Oxford Dictionary (oxforddictionaries.com), over 250.000 words? Write & Improve indirectly answers this question when addressing the issue of native speakers testing their software:

Computers are not yet capable of understanding a piece of text in the same way as a human being. They do not have the same context or life experience that a human can bring to bear.

In other words, computers may be objective but they are not yet sophisticated enough to replace humans, at least in this field. That is why the software relies on statistical analysis of a large number of features” that are matched with “the same features extracted from a large corpus of ‘training data’”. This “training data” they mention are essays from actual EFL students from which the level of proficiency in the English language has already been measured.

These features act as proxies for the student's level of attainment: some are indicative of good writing, others of poor writing. Write & Improve combines these positive and negative indications together to generate the final score for a piece of writing. (…) Given writing created by genuine EFL students (the kind of writing that Write & Improve has been trained upon) you will see very accurate results.

(writeandimprove [www])
Although they recognize the issues that this essay has discussed and will discuss, they claim to have found accuracy and validity through the size and nature of their extensive database.

**Figure 2.** What Write & Improve interface can look like (Write & Improve [www])

If one were to get back to basics and what language testing is about, and was about before video killed a radio star; what kind of competences does one look at? What is the difference between knowing and expressing?

Canale and Swain (1980) express the difference between communicative competence and communicative performance as something that is vital for language testing (Fulcher & Davidson 2007:38-40). Communicative competence includes grammatical competence, sociolinguistic knowledge and strategic competence while communicative performance simply implies demonstrating your knowledge in a performance (2007:38). Ideally, for language testing, “tests should contain tasks that require actual performance as well as tasks or item types that measure knowledge” (Fulcher & Davidson 2007:39).

A written text in English inevitably demonstrates said communicative performance. However it does not apply the second criteria of measurability very well, that is: tasks or items. This relates to the problematic aspect of human unreliability. This aspect is applicable to language testing of the written national test in English, that relies entirely on the varying ability of teachers to assess communicative performance. Any rating system that relies on human input is naturally vulnerable to the subjectivity of the
individual. Such an issue has been minimized in more defined environments where the rating system and the individuals carrying out the system are internal – they are trained and active in the same place. “There must be a group of people whose ability to place language models into categories has evolved over time, and into which newcomers can be socialized” (Fulcher & Davidson 2007:96-7) and become “adept” as Lowe (1986) boiled it down to. Lowe (1986:392) found that new teachers when starting to assess proficiency with the help of guidelines typically would focus on separate sentences of said guidelines and isolate different criteria in a problematic way. They would ask themselves: does the pupil aptly use adjectives? Instead of asking: Does the pupil consistently show a certain language proficiency? Lowe developed his reasoning:

One may say, of course, that the Guidelines reflect this process orientation less well than they might. But one document cannot be all things to all people – to test designers, to raters, to course developers, to materials writers, to classroom teachers, and to administrators. The Guidelines must fail in many of these demands because words do not always capture the essence of a concept and because the Guidelines were originally designed to outline, not to describe the system exhaustively; they function more as a constitution than a Napoleonic code. . . The Guidelines’ greatest utility may lie in their use as a framework . . . Lowe (1986:392)

He found guidelines, such as the guidelines that the model texts come from, to be important but that they required something more from the examiner. He found it more important when a pupil could provide a sufficiently long text that showed “sustained creativity and generativity”(Lowe 1986:393) and less important to show this in “bits and pieces of the language” (393).

While Erickson (2009:20) has found that assessing and testing language can be done with both an analytical approach and a holistic approach, although she found holistic to be favorable, Lowe (1986:392-5) advocated more strongly for a holistic approach and claimed that the way to master such an approach was not just to use the guidelines but also to interpret them with insight – a challenge for those that have not yet become adept. He found that the ability to assess correctly was strongly connected to long lines of tradition in assessment, of interpretation being passed down to novices (392). Although Lowe did not define the process of becoming adept to be time-sensitive, one can relate this challenge of adeptness to both work experience (see page 12) and the fact that there is a steep learning curve within the teaching profession. Teachers dramatically increase their ability to assess reliably in the first two to three
years before stagnation occurs and they, interestingly enough, nearly stop improving (Sundqvist 2017:10).

Another core challenge, of course, is validity. The term itself is used broadly across many fields but within education and assessment it is considered absolutely vital (Newton 2013: lecture) (Fulcher & Davidson 2007:3-12). In the context of education validity can be understood as a means of discovering whether a test “measures accurately what is intended to measure” (Hughes 1989: 22).

Validity is, however, a rather abstract notion as it encompasses many different usages and models (Fulcher & Davidson 2007:4-11) but essentially it is about how valid, accurate or well represented something is (Newton 2013). In this paper, the importance of validity will not focus on the test format from the student’s perspective, such as how the student perceived the instruction or whether the writing assignment for the national test in Norway and/or Sweden is inclusive of varying social groups. Such considerations fall within the category of language testing but not the current perspective. Instead, focus will be on the validity of the model texts that the teachers and sensors are to follow and, more specifically, are supposed to validly represent grades. To determine whether they are valid in regard to the European set standards that, while non-obligatory, are prioritized reference points for both countries, is the task at hand. Such a task is challenging but possible due to the CEFR’s work with creating a framework for languages within Europe. The English section of this framework was produced through eight stages in collaboration with the University of Cambridge. Commissioning, pre-editing, editing, pre-testing/trialling, pre-test review, paper construction, examination review and question paper construction (QPP) are the eight stages produced in their work with language testing for English (Council of Europe 2005). From this foundation sprung the software Write & Improve a few years later (see Method section) which should ideally represent the CEFR proficiency levels from A1 to C2. The assumed complexity of language testing from a historical perspective might thus be challenged by digital innovations like Write & Improve that standardize and can evaluate more objectively. In the light of such innovative thinking it can be important to remind ourselves that, as for now, this type of test is not assessed by any calculation or robot. It is assessed by professionals that are vital to the process, Sundqvist emphasizes, considering the Swedish teacher plays a three-fold role in carrying out the national test: teacher, test administrator and examiner (Sundqvist 2017:7) It could be argued that the Norwegian counterpart would then be playing a four-fold role as a sensor with a more
active role in the construction of the test, as well. These teachers or sensors are then prone to extenuating factors that can affect how they approach the assessment of the national test (Sundqvist 2017:9) and by default also how they approach the guidelines for said test.

Certification, that is whether the examiner has a licence to do so or not, has been proven by both Sundqvist’s research and others before her to have next to no impact (2017:7, 29) Work experience, on the other hand, has been proven by multiple studies to have a significant effect on assessment in general (10). Furthermore, experienced teachers more often than unexperienced ones assessed in alignment with test administration. They also tended to make faster and more functional decisions (33).

The Swedish Schools Inspectorate also found that students from certain groupings such as upper-class girls who were expected to perform well tended to be given more generous grades than those given by an objective second examiner. The same goes for groups that are expected to perform poorly such as working class boys, receiving poorer grades than they ought to (Swedish Schools Inspectorate 2013:19-25). Evidently, the issues of subjectivity contra objectivity in language testing are many and this along with the differences that do exist in “[t]eacher practices and views regarding the test” (Sundqvist 2017:36) teachers become problematic examiners. Let it be clarified, however, that this paper will not take such arguments further than this reached conclusion.

3 Method and Material

3.1 Material

The most important materials are the model texts that are provided to teachers when they are assessing the national test in English writing. The Norwegian document is slightly less detailed due to the sensors having had special training before going into the assessment process. The Swedish one is more detailed with a general introduction on how it was developed and a larger grading chart that was not considered necessary in Norway since it was covered in their training (Interview Matilda Burman).

The tool Write & Improve and the database EVP are also important, to cover the CEFR-aspect of this paper’s aims and compare the quality of language between the previously mentioned model texts.
The Norwegian model texts

The Norwegian model texts (see 4.1) and guidelines are from the spring of 2014 and feature texts about international role models such as Martin Luther King and Malala Yousafzai. The instructions for the Norwegian model texts are as follow:

Assignment instructions

Task 1
In the preparation material you have read a newspaper article about Nelson Mandela. Answer the following:
• Why is he a role model for so many people in the world?
• Who do you look to as a similar role model and why?

Task 2
In Appendix 1, you will find an extract from Dr. Martin Luther King, Jr.’s I Have a Dream speech.
Read the extract and answer the following:
• What is Dr. Martin Luther King, Jr.’s main message in this speech?
• How does the language he uses strengthen the message?

Task 3A
Not everyone can have such an impact on the world as some of the people you read about in the preparation material. Change can also occur through small steps.
Create a text in which you talk about small steps you could take to make a change.
Choose a suitable title and type of text.

Task 3B
Many of the texts in your preparation material have been about people being stereotyped and prejudiced against.
Choose at least two people from the material you have worked with and discuss the following:
• why they were stereotyped or prejudiced against
• what happened because of this
• what decisions they took
• the result of their decisions

Task 3C
In your preparation material you have read about people who have overcome or are living in difficult life situations.
Compare a character from your preparation material with another character from your English course and discuss how they deal with difficult life situations. Your text should include:
• a brief comparison of the characters
• a clear description of the two situations
• a discussion of what the characters do to overcome the difficult situation
• the consequences of their actions

Task 3D
Look at the Norman Rockwell painting on the title page. The title of the painting is “Moving Day”, and it is from 1967. Create a text inspired by the painting. Include the following:
• describe the painting and its setting
• choose one of the children in the painting and describe what he or she is thinking about
• discuss what the painting reveals about race issues in the USA
(Guidelines Nor 2014:2)

The pupils were expected to complete task 1 and 2 and afforded a choice between 3A, 3B, 3C and 3D as for their final task. The first text to be presented from the model of Norwegian model texts will be called N2. “N” for Norway and “2” for the grade it represents. The following texts will follow the same pattern for ease of reference, from N2 to N6. Grade 3 has two representative model texts, the first will be N3i and the second N3ii. Additionally, most of the model texts are longer than 600 words and have been split up for the Results section.

Table 1. Overview of the Norwegian model texts

<table>
<thead>
<tr>
<th>Model texts</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2</td>
<td>378</td>
</tr>
<tr>
<td>N3i</td>
<td>705</td>
</tr>
<tr>
<td>N3ii</td>
<td>893</td>
</tr>
<tr>
<td>N4</td>
<td>1346</td>
</tr>
<tr>
<td>N5</td>
<td>1230</td>
</tr>
<tr>
<td>N6</td>
<td>1367</td>
</tr>
</tbody>
</table>

Table 1 gives a basic overview of the different Norwegian model texts. The model texts are written by pupils in the 10th grade (10:e trinn, 16 year olds) and the assessment of these are presented as exemplary in the document (NDET 2014). It is important to note that both the age of the pupils and the purpose of the published texts match the Swedish counterpart. The fact that the Swedish pupils are in the 9th grade and the Norwegians in
the 10th might appear contradictory to that statement. However, the first year in Swedish school system is called förskoleklass (pre-school) (Skolverket, förskoleklass [www]). Consequently, year 9 in Sweden and year 10 in Norway are effectively pedagogically equivalent and comparative.

The Swedish model texts
The Swedish model texts (see 4.2) and guidelines are from the spring of 2013 and the assignment is called Our Time – My Story. In it the pupils are asked to write a contemporary autobiographical text (with simpler words). The instructions are divided into three: 1. Describe your life right now, 2. Explain how different styles and trends influence you now, 3. Discuss one or two issues that are important to you or to other people, today and in the future. (SNAE Ämnesprov 2013) The test is for pupils in the 9th grade and the model texts are, just like for the Norwegian ones, presented as exemplary in the document. The instructions given to the pupils that wrote the model texts are as follows:

At LifeStory we are making an online collection of texts written by people all over the world. Our idea is to create an archive of texts about what living in the 21st century is like. We believe that the stories will help future generations to understand our time. You are invited to write a text for the LifeStory Archive – about yourself and the time you live in.

Plan your writing and make sure that you have time to write about all three parts (A-C). Altogether you should write between 250 and 500 words. Use the following points:

A. Describe your life right now,

B. Explain how different styles and trends influence you now – or have influenced you before. It could be in, for example, music, clothes or technology.

C. Discuss one or two issues that are important to you or to other people, today and in the future. It could be about health, the environment, politics, religion, etc.

(SNAE Ämnesprov 2013)

The pupils were expected to write about themselves and the society around them following these three points. The Swedish grading system goes from F – A, F being a fail, E the lowest pass, and A the top grade. In between D (overwhelming C level but some criteria at E), C and B (overwhelming A level but some criteria at C). As for the model texts there are eight Swedish ones, neatly summarized in Table 2.
Table 2. Overview of the Swedish model texts

<table>
<thead>
<tr>
<th>Model texts</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>S = Swedish</td>
<td></td>
</tr>
<tr>
<td>E,D,C,B,A = grade</td>
<td></td>
</tr>
<tr>
<td>SE-</td>
<td>449</td>
</tr>
<tr>
<td>SE+</td>
<td>222</td>
</tr>
<tr>
<td>SD-</td>
<td>306</td>
</tr>
<tr>
<td>SD+</td>
<td>408</td>
</tr>
<tr>
<td>SC-</td>
<td>384</td>
</tr>
<tr>
<td>SC+</td>
<td>503</td>
</tr>
<tr>
<td>SB</td>
<td>371</td>
</tr>
<tr>
<td>SA</td>
<td>431</td>
</tr>
</tbody>
</table>

EVP – the database

What kind of material would then be needed to get a better understanding of the CEFR? To help teachers understand what the CEFR means for the English language specifically, a website and database was developed by Cambridge University, on behalf of the Council of Europe (englishprofile.org). It is called English Profile. English Profile has done much to categorize language in accordance with CEFR, describing “what aspects of English are typically learned at each CEFR-level” (englishprofile [www]). The English Profile material this study will use is EVP, the English Vocabulary Profile, that grades all words and uses of them from A1 to C2 to help us decipher what is considered advanced level and what is not.

On a rudimentary level A, B and C can be divided into basic, intermediate and advanced; according to a guide published in connection to the English Profile called “Introductory Guide to the Common European Framework of Reference (CEFR) for English Language Teachers” (Cambridge University Press: 2013:2) The English profile also set out to concretize the proficiency levels for CEFR English using a vocabulary profile (EVP).

The EVP is meant to illustrate how different levels or proficiencies of the English language can be observed through different usage of words and syntax. For example “fine”, the adjective, is considered something new to an A2 learner while “fine”, the noun, is tied to a B1 level. These are two different words spelled the same but one and the same word can also be used in different ways for different levels of proficiency. To
“come from” (A1) is considered a more basic use of the word “come” than a phrase like “come in” (to enter)(A2). (English Vocabulary Profile [www])

Evidently, the English Profile attempts to clarify how the English language interacts with the CEFR on a general level but more specifically for the model texts that this study is about, the Council of Europe supplies an overview of the expectant language levels for written production.

Table 3 – Overview of CEFR-grades, written production

<table>
<thead>
<tr>
<th>OVERALL WRITTEN PRODUCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C2</strong></td>
</tr>
<tr>
<td>Can write clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader to find significant points.</td>
</tr>
<tr>
<td><strong>C1</strong></td>
</tr>
<tr>
<td>Can write clear, well-structured texts of complex subjects, underlining the relevant salient issues, expanding and supporting points of view at some length with subsidiary points, reasons and relevant examples, and rounding off with an appropriate conclusion.</td>
</tr>
<tr>
<td><strong>B2</strong></td>
</tr>
<tr>
<td>Can write clear, detailed texts on a variety of subjects related to his field of interest, synthesising and evaluating information and arguments from a number of sources.</td>
</tr>
<tr>
<td><strong>B1</strong></td>
</tr>
<tr>
<td>Can write straightforward connected texts on a range of familiar subjects within his field of interest, by linking a series of shorter discrete elements into a linear sequence.</td>
</tr>
<tr>
<td><strong>A2</strong></td>
</tr>
<tr>
<td>Can write a series of simple phrases and sentences linked with simple connectors like “and”, “but” and “because”.</td>
</tr>
<tr>
<td><strong>A1</strong></td>
</tr>
<tr>
<td>Can write simple isolated phrases and sentences.</td>
</tr>
</tbody>
</table>

Note: The descriptors on this scale and on the two sub-scales which follow (Creative Writing; Reports & Essays) subscale have not been empirically calibrated with the measurement model. The descriptors for these three scales have therefore been created by recombining elements of descriptors from other scales. (CoE, Structured Overview: 23)

Looking at these descriptions and understanding them is different from actually being able to identify and discern the qualities sought after. Furthermore, even if one did identify a segment of a produced written text as “expanding and supporting points of view at some length with subsidiary points” it can be challenging to take on a holistic approach for a system like the CEFR - especially for someone used to their national non-comparative language levels. For this reason a software like Write & Improve, the problems highlighted in section 3 withstanding, can be useful and insightful. There one can get a sense of what level of CEFR proficiency a text is at from the leading authority on the matter, Cambridge University.
Write & Improve

Write & Improve is a software developed by Cambridge University on behalf of the European Union, to create a tool in synchronization with the CEFR. It is meant to help English language learners improve their writing and uses detailed calculations to put such texts into a certain CEFR proficiency level. Laurie Harrison (2017) explains the technical function of the software of it in a featured interview at eltJAM.

She describes a “supervised machine learning . . . that is fed training data” from the Cambridge Learner Corpus, a database with over 30 million error-annotated words. The data is from L2 writers, learners that are not native speakers of English, and Write & Improve and its algorithms are designed to learn and “speak learner English”, meaning it is smart and adapts to data input. Furthermore, Harrison explains that while it excels at evaluating L2 English, proper English or “purple prose” from a native speaker will “confuse it enormously”, she says (eltJAM [www]).

Write & Improve is, thus, an intuitive and continuously improving but not necessarily perfect software. It will accurately detect problematic language in English written by L2 learners, i.e. the type of language used in the model texts. To help the pupil improve (which is the primary purpose although its results are supposed to be accurate enough for language testing), the software supplies markings and highlights parts of the text; white-yellow highlighting means the text makes sense but the grammar or syntax could be improved, full yellow highlighting means there is a greater need for improvement in more ways than one, arrows mean that there is a missing word, stars with a questionmark in them mean suspicious word: implying the software is not sure, but suspects that word is not the intended one, and, finally, exclamation mark means noteworthy spelling mistake - usually the most commonly used words from my interpretation (this, that, until etc.), as it chooses to use this in some cases over others. These annotations are straightforward and rather easy to understand, although they (the annotations) do leave much to desire in terms of deeper language analysis. For example, there is no general feedback on what to improve (pronouns, syntax etc.) and the specific feedback (arrows, exclamation marks etc.) does not suggest specific improvements or any detail beyond the nature of the mistake. This makes the analysis and results less clear.

Despite that, Write & Improve appears to be the best available tool for this study, considering the lack of alternatives and the fact that it was developed for the specific purpose of including CEFR into a software that can be used and understood in different
countries. As a consequence, it is predicted to work well for a comparative study such as this, albeit with more focus on its result and less on its language analysis.

3.2 Method

The method of this paper is ambitious in the sense that it hopes to help achieve an analysis that can state whether or not the model texts live up to the expected standard, in other words to discern quality in language. This is not a straightforward task and so it will be tackled in more than one way for as broad an approach as possible, but within the confines of this type of research paper. First, the paper will give way for digitalization and use the software Write & Improve to evaluate the texts holistically. Secondly, the database EVP will be put to use and I will pick the model texts apart word by word, more analytically determining how advanced a vocabulary the model texts can boast with. Thirdly, I will tie in my own observations of the context (2.1 and 2.2), the secondary sources (see section 2.3), the model texts and the results from Write & Improve and the EVP analysis (see section 4), combining the facts into an analysis (see section 5) which leads to a conclusion (see section 6). This database will be used thusly in section 4 (results), to attempt understanding the qualities of the different model texts.

The selection of which texts should be tested, and which should not, makes for an interesting methodological predicament as the outcome of them might prove useful to varying extents. If a Swedish model text that is meant to represent the grade C and a Norwegian one meant to represent the grade 3 (from 2 to 6) are both analyzed – are they comparable? Instinctively one might say yes, but with different systems come different connotations and it might not be as straightforward as that. However, the differing grade systems and the fact that the spark (Bruun & Diaz’ article) was concerning the lowest requirements for a pass, both lead to the conclusion that focus ought to be on the lowest graded ones. Nevertheless, for the sake of validity and full disclosure, all model texts will be submitted into Write & Improve to see if there are any x-factors or interesting observations from the remainder of the model texts that might contribute to the discussion.

An additional methodological aspect to consider is the reliability of each result. To clarify with an example, model text SB is the Swedish model text meant to represent the grade B (more on this in Results and in Attachment 1) and it contains some anonymized words. The location, school and family members that the pupil described in his or her text where replaced with NN, XX etc. for privacy. When I experimented and
changed those anonymizations into more fitting words such as London, London school and Linda, the grade improved from B2 to C1. In this case, those anonymizations affected the over-all CEFR grade Write & Improve supplied. Likely, the text was on the verge of being considered C1 and those small tweaks pushed it across the finishing line. Whether that should be taken into account for the raw findings is an interesting but difficult question. Would teachers reading these model texts let the anonymizations affect their judgement? I find that unlikely but nevertheless, this study will respect the original phrasing of the model texts and include anonymizations or other exceptions in its raw findings, but rather discuss the impact such small but sensible corrections can have in the analysis of the computer-generated results. Every word counts and sends a message to the test examiners that use the model texts as standards. That is why the results from EVP will be relevant and why that database is an important component of this study. SE- and N2, the lowest pass model texts, will be manually picked apart word by word and cross-checked with the EVP-database to give an idea of the advancedness of the vocabulary. In theory, a CEFR-aware test examiner that notices a C1-word in model text SE-, would not let that affect his or her grading. If there’s a pattern of more advanced vocabulary, however, it should raise the bar for getting an E in Swedish national tests.

For the Write & Improve results, reservations must be made for the three-part-assessment structure in the guidelines following the Norwegian model texts. Content, text structure and language proficiency are brought up as main things that were considered in the assessment of the model texts which means that content is regarded equal in importance to text structure and language proficiency (Guidelines Nor 2014:2-3). This is something that Norwegian scholars have discussed as a potential problem, referring to the washback-effect, according to Hellekjær (2011:42-7) and Ellingsund (2009) found in her interview-based study that teachers found it hard to focus on increasing their students proficiency as content took up too much time (Ellingsund 2009:79). In comparison, the Swedish guidelines express that the most important thing to look for are pupils’ ability to formulate and communicate in written English production. Whether the pupil stays on point regarding the assignment or if the content of the text makes sense is important for the goal to adapt to purpose, recipient and situation – but over-all it is considered secondary to the linguistic quality in importance (Guidelines Swe 2013: 3). As for Write & Improve, it can readily account for text structure and language but has not been instructed to look for content, although it has
that feature too. Users can choose to add assignments and criteria in terms of content when using the software for this purpose. However, the credibility of such a feature in the analysis could prove questionable. Will instructions fed to Write & Improve match the original assignment given to Norwegian pupils? And, more importantly, will its understanding of those instructions match that of Norwegian sensors or Swedish teachers? For these reasons the only criteria fed to the software were for the minimum amount of words to be 1 and maximum 600. 600 words is a maximum set by Write & Improve as it made for such long texts (writeandimprove.com), some of the examples that are going to be examined extend beyond that 600-word limit but are neatly separated into different tasks and can be assessed all the same. As for the assignment: “Write about anything” was given as instruction, liberating the software of that element and giving full focus onto language. With these things in mind, model texts (that are meant to represent a certain grade) from both Sweden and Norway will be tested by this program, i.e. copied and pasted into the software’s writing field.

Although the countries have the same numbers of passing grades available, that is the 5 grades that qualify as a pass (E-A and 2-6), the Swedish guidelines encourage teachers to use “weak” and “strong” Cs for example, often simply referred to as a C- or a C+. As this logic is applied to all grades the scale becomes doubled, from 5 to 10 passing grades; although in actuality only 9 texts are presented in the document. Strong F (fail), weak E, strong E, weak D, strong D, weak C, strong C, B and A. (Guidelines Swe 2013:3). This is another issue to consider when comparing the two countries’ approach to model texts.

4 Results

The results section of this paper is geared towards showcasing relevant and countable data to support the broader analysis of the model texts.

4.1 Write & Improve

This section contains three tables showing the Write & Improve software’s over-arching take on the advancedness of the model texts. These are presented with the code, the given CEFR grade, the number of words and other potentially useful information.
Table 4 – Write & Improve results, Norwegian model texts

<table>
<thead>
<tr>
<th>The Norwegian model texts</th>
<th>Norwegian grade</th>
<th>CEFR grade</th>
<th>Number of words</th>
<th>Divided parts for enabling W&amp;I-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2</td>
<td>2</td>
<td>B1</td>
<td>378</td>
<td></td>
</tr>
<tr>
<td>N3i</td>
<td>3</td>
<td>~B2</td>
<td>705</td>
<td>a) C1 b) C2</td>
</tr>
<tr>
<td>N3ii</td>
<td>3</td>
<td>~B1</td>
<td>893</td>
<td>a) B2 b) C2</td>
</tr>
<tr>
<td>N4</td>
<td>4</td>
<td>~C2</td>
<td>1346</td>
<td>a) C1 b) C2 c) C2</td>
</tr>
<tr>
<td>N5</td>
<td>5</td>
<td>~C1</td>
<td>1230</td>
<td>a) C1 b) B2 c) C2</td>
</tr>
<tr>
<td>N6</td>
<td>6</td>
<td>C2</td>
<td>1367</td>
<td>a) C2 b) C2 c) C2</td>
</tr>
</tbody>
</table>

The CEFR grades that feature a “~” in Table 4, like N3i for example, imply that the separately evaluated “parts” have yielded different results. As a consequence an average grade has been awarded with those parts in mind. This is for practical purposes, however the differing results is of course an interesting observation that will be discussed further in the Analysis section of this paper.

Table 5 - Write & Improve results, Swedish model texts

<table>
<thead>
<tr>
<th>The Swedish model texts</th>
<th>Swedish grade</th>
<th>CEFR grade</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE-</td>
<td>E</td>
<td>A2</td>
<td>449</td>
</tr>
<tr>
<td>SE+</td>
<td>E</td>
<td>A2</td>
<td>222</td>
</tr>
<tr>
<td>SD-</td>
<td>D</td>
<td>A2</td>
<td>306</td>
</tr>
<tr>
<td>SD+</td>
<td>D</td>
<td>B1</td>
<td>408</td>
</tr>
<tr>
<td>SC-</td>
<td>C</td>
<td>B2</td>
<td>384</td>
</tr>
<tr>
<td>SC+</td>
<td>C</td>
<td>C2</td>
<td>503</td>
</tr>
</tbody>
</table>
Table 6 - Write & Improve results, Norwegian model texts

Blue statistics are **Swedish**, red are **Norwegian**.

<table>
<thead>
<tr>
<th>Model texts national grade</th>
<th>Model texts CEFR grade</th>
<th>Words</th>
<th>Model texts national grade</th>
<th>Model texts CEFR grade</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-</td>
<td>A2</td>
<td>449</td>
<td>2</td>
<td>B1</td>
<td>378</td>
</tr>
<tr>
<td>E+</td>
<td>A2</td>
<td>222</td>
<td>3i</td>
<td>~B2</td>
<td>893</td>
</tr>
<tr>
<td>D-</td>
<td>A2</td>
<td>306</td>
<td>3ii</td>
<td>~B1</td>
<td>1346</td>
</tr>
<tr>
<td>D+</td>
<td>B1</td>
<td>408</td>
<td>4</td>
<td>~C2</td>
<td>1230</td>
</tr>
<tr>
<td>C-</td>
<td>B2</td>
<td>384</td>
<td>5</td>
<td>~C1</td>
<td>1230</td>
</tr>
<tr>
<td>C+</td>
<td>C2</td>
<td>503</td>
<td></td>
<td></td>
<td>1367</td>
</tr>
<tr>
<td>B</td>
<td>B2</td>
<td>371</td>
<td>6</td>
<td>C2</td>
<td>1378</td>
</tr>
</tbody>
</table>

Although it may be true that Swedish E- and E+ texts are not necessarily meant to match Norwegian 2 texts, for comparative purposes and for creating an overview, the two systems have been placed parallel to one another in Table 6.

4.2 Further examining of N2 and SE-
This section will show how the markings of Write & Improve come across and what raw data can be found on the analysis of N2 and SE—specifically, the two model texts deemed most relevant to analyze. Additionally, I have looked at all the words in both texts individually and checked them against the EVP (see 3.1 Material).

**Screenshots of the results for N2 and SE—**

Figure 8. The result of checking N2 against Write & Improve.

Figure 9. The result of checking SE—against Write & Improve.

Abbreviations and markings to be used in this section and the Analysis:

N2 = Norwegian model text with grade 2 (lowest)

SE = Swedish model text with grade E- (lowest)

t1, t2, t3 = task 1, task 2, task 3.

A1, A2, B1, B2, C1, C2 = CEFR levels of proficiency (from lowest to highest)
**Figure 10.** Markings as explained at writeandimprove.com

**N2** has 684 words in total and 13 of these were suspected for not being entirely correct, such as for example the second use of “the” in “she thinking the women had the powerful the men”. This use is grammatically incorrect and thus correctly acknowledged. The 13 words amounted to 1.9 % of the entire text. Only 2 instances of something missing *before* and 3 instances of something missing *after* were identified. In other words, 0.7 % of the entire text had a missing word. Also, there were 17 misspellings noted, making up 2.5 % of the text.

Specific word-markings were relatively few and in total only made up 5.1 % of the model text, instead much of it was marked as faulty sentences. There were a total of 356 words in the sentences that had some problems (52 % of the model texts) and there were a total of 95 words in the sentences that could be improved (13.9 % of the model texts).

**SE** has 225 words in total and 3 of these (1.3 %) were suspected for not being entirely correct. A consistent example of this would be “im” which is intended to read “I’m” but the software is not buying it. There was 1 case of a word missing something after it, making up 0.4 % of the total text and when it comes to spelling, 7.1 % of the words in the Swedish model text were marked as misspelled. That is 17, which is the same as N2 had, but notably a higher percentage of mistakes since the Swedish model text is that much shorter.

Regarding marked sentences, they amassed to almost the entire text with 180 words in sentences with some problems (80 % of the model text) and 8 words in sentences that could be improved in some way (3.5 % of the model text).
4.3 EVP results

The texts have here been manually checked against the EVP in terms of how advanced the vocabulary is from A1 to C2. Since most common words are A1 it is not worth noting them down. Consequently, all the words marked with their respective proficiency level tag are from A2 to C2.

**Vocabulary of N2**

<table>
<thead>
<tr>
<th>A2</th>
<th>B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Century</td>
<td>Struggled</td>
</tr>
<tr>
<td>Among</td>
<td>Stand for (represent)</td>
</tr>
<tr>
<td>Prize</td>
<td>Against (a cause)</td>
</tr>
<tr>
<td>Several</td>
<td>Peace</td>
</tr>
<tr>
<td>Going to (do something)</td>
<td>Rights (Human)</td>
</tr>
<tr>
<td>Receive</td>
<td>Speech</td>
</tr>
<tr>
<td>International</td>
<td>Draw attention</td>
</tr>
<tr>
<td>While</td>
<td>Nation</td>
</tr>
<tr>
<td>Himself</td>
<td>Slaves</td>
</tr>
<tr>
<td></td>
<td>Equality</td>
</tr>
<tr>
<td></td>
<td>Major</td>
</tr>
<tr>
<td></td>
<td>Spread</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Elected</td>
</tr>
</tbody>
</table>

**Vocabulary of SE-**

<table>
<thead>
<tr>
<th>A2</th>
<th>B1</th>
<th>B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hobby</td>
<td>Trends – B1</td>
<td>Make money (earn) – B2</td>
</tr>
<tr>
<td>Nobody</td>
<td>Technology – B1</td>
<td></td>
</tr>
<tr>
<td>Fail</td>
<td>Screams – B1</td>
<td></td>
</tr>
<tr>
<td>Interested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Way</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grow (up)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Comparison on the number of advanced words in SE- and N2

<table>
<thead>
<tr>
<th></th>
<th>SE-</th>
<th>N2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>B1</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>B2</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>C1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>C2</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 7 shows a comparison on the number of advanced words found in SE- and N2. N2 very clearly boasts a more advanced vocabulary, even proportionally, when one considers the number of total words in the respective texts.

5 Analysis and Discussion

5.1 N2 and SE-
This discussion will be about the results in general but will largely focus on SE- and N2, this because of the lowest passing grades being more easily comparable between two different systems of marks. An argument could be made for focusing on SA and N6 on the other end of the spectrum but considering the Bruun and Diaz’ article (2017), the concrete reference point that the official expected proficiency level of B1 in Sweden makes for and the more dramatic difference between a pass and a fail, rather than a B or an A, SE- and N2 are far more convenient, fruitful and vital. Norwegian texts are fragmented and in total longer than the Swedish ones are. They exhibit bigger vocabulary but how much of that is due to pure quantity and how much is due to the nature of the exam? One must consider both. When it comes to the nature of the exam, the premise of the Norwegian task 1, 2 and 3 is that they read a small text on the subject and reflect based on that. Some passages in N2 are blatantly obvious copied from this information such as the quotes in task 3. These are naturally overlooked both by me analyzing them and most likely also the examiners and the authors of the document (that picked out the model texts) but not by Write & Improve, according to my tests, removing and adding the extraordinary segments. There is a grey zone, however, between blatantly obvious copied vocabulary (He issued the memorable Emancipation proclamation in 1863) and completely independent vocabulary (they thier time and their life and save most of others), that is hard to decipher.

Take for example this quote from t3: “He **answered** famous also in that
divided of population”. Although the syntax and the spelling is far too off to be blandly copied, it looks to say something like “He famously answered [hatred] within a divided population” which is perhaps too complex language to be independently concocted. To demonstrate, the adjective divided is considered a B2 word according to EVP, which is higher than the over-all proficiency level as deemed by Write & Improve. Even so, using high-proficiency vocabulary in a poor manner might actually benefit the text in the eyes of test-creators and examiners. By this I mean that it shows an attempt at contextualizing and using a word by one’s own means instead of, as in the case of the ‘emancipation’ quote, just copying. In other words, one can speculate that a single outstanding impeccable phrase would be disregarded but a less perfect phrase using high-end vocabulary would not.

The model text N2 contains 378 words and was supposed to represent the grade 2 (from 1-6) and a pass in Norway. It was deemed B1 level in CEFR proficiency by Write & Improve. Here, we can see at least one sentence out of the ordinary where at the last paragraph the pupil quotes Abraham Lincoln saying “Whenever I hear anyone arguing for slavery, I feel a strong impulse to see it tried on him personally”. The quote from Lincoln along with the quote about the issuing of the “emancipation proclamation” could possibly, although unlikely, have been produced by the pupil him-/herself. What matters in this analysis, however, is the sensors’ perspective on the text as it is presented to them and considering the quotes’ vast difference in quality to the rest of the text they would likely be disregarded as verbatim quotes in an assessment based on validity.

This ability to omit exceptions might not be something the Write & Improve software is equipped to do, as discussed in section 3 (Method and Material). The quotes might thus be undesirably overbearing in the software’s calculations, creating an unjust result. I had them omitted, however, and the text tested again, although no change in proficiency level occurred. The Norwegian test-creators also noticed this discrepancy as they noted that N2 showed “signs that different parts of the text are haphazardly put together to fit the demands of the task” (see appendix B).

This “human” perspective on the texts has been applied to make the comparison between NDET/SNAE assessments and Write & Improve’s CEFR-assessment as valid as possible. t1 and t2 make up roughly half of the model text and result in the grade A2 which notably is below the expected level of language at that age in Sweden. When it comes to age and expected proficiency there are, as mentioned, as of yet no such charts in Norway. For this reason, the analysis will have its focal point in the Swedish
expectancy. t3, which by some regard, is a stronger text than t1 and t2 grade at B1 which matches the expected level of proficiency.

The Norwegian texts are by design chosen and evaluated through three categories of criteria: content, text structure and language. It should be clarified that text structure is not simply larger structuring of the text into paragraphs but also about cohesion and syntax. To each of these criteria there are response or comments that in relation to the criteria remark upon the quality of the model texts.

When it comes to content, it is criticized for being patchwork but one key expression to take away from it is “certain degree”, as in “it fulfills the criteria to a certain degree”. The same can be said for the text structure but not for language where “small degree” and “some degree” is used. These phrasings are perhaps not the only things to take away from the summarized responses but they are revealing.

If the weakest of the given responses were in relation to language and N2 still exhibits a higher English proficiency than SE-, it might mean that the gap in quality between these texts is even bigger than it seemed at first glance.

It is necessary to take into account that the Swedish exam expresses a greater focus on language and sidesteps the importance of content (at least it does so in the introduction). In general, SE- is poorer in terms of vocabulary and varied expressions but remains comprehensive through and through while N2 is less reliable and at times completely incomprehensive. See, for example; …, *man how to solve for the strict things out of lost the people and he is going to die about many to give his life for all* (N2, Udir.no) This type of differences could possibly be related to Norway not fully having implemented the ideas from the CEFR. When looking more strictly at the Write & Improve analysis one can observe patterns and “count” the quality of the text.

5.2 Comparison of results
Quantifying the Write & Improve results in this way is not meant to replace its calculations but a means of understanding them. When considering sentences marked as problematic (see 4.2), the 80 % on SE- is remarkably high. When comparing these two quite different texts that both reflect the lowest passing grades in Sweden and Norway, one needs to keep in mind the already mentioned flaws of the software and also the details listed below. And with those in mind, one might say that the high percentage of solidly colored words in SE- could simply be happenstance due to its brevity in comparison to the long Norwegian text. 225 words could easily become both 80 % and
0 % since it is a very short text. That line of reasoning may seem sound, but neglects the perhaps obvious but not too overstated fact that the texts are not picked at random. The model texts are specifically selected to be representative of the language proficiency pupils need to show in order to achieve a passing grade, such circumstantial arguments can be kept in mind but do not render the potential importance of statistics moot. On the other hand, one marking of forgot-sth-after at 0.4 % is likely not particularly relevant to explore further. When looking at specified markings by Write & Improve (i.e. not the broader markings that stretch to more than one word) we can conclude the total percentage for all of these land on 5.1 % (N2) and 8.7 % (SE-) respectively, a difference not as great as the underlined one but nonetheless noteworthy and implying that the Norwegian text attains higher quality overall and not just when it comes to t3.

Even more noteworthy is the significant difference in the number of suspicious words, where N2 has more than SE-, maybe due to Norwegian pupils being fed words they do not know how to use. As SE- is shorter the difference in percentage when comparing between the two texts is not as significant as comparing the internal difference; the balance between suspicious words and incorrect ones for both texts. In N2 it is quite balanced while in SE- the incorrect words are almost six times as many as the suspicious ones, implying an overbearing issue with spelling that one can confirm when looking more closely at the text. In doing so, one can observe a lot of strange looking words that are not marked with spelling. This is not particularly surprising since words like “inportent” are so far off the originally intended word “important”, the software likely does not recognize it is an issue with spelling since it can not identify what it should actually say. Perhaps, if the pupil had written “importent” with an M, it would recognize the specific error and trigger a 17th spelling marking (in addition to the 16 already mentioned in 4.2).

Additionally, the spelling of the words naturally affects the quality of the text. Sometimes the degree of intelligibility affects the software’s perceived quality of the text as well, as was the case with ‘inportent’. Whether it is ‘divided’, ‘emancipation’ or ‘inportent’ it is clear that the vocabulary plays a big role and does so in both countries’ curriculum as well. Consequently it would be remiss not to go into more detail there, especially with the EVP at hand.

As was the case with the preliminary results for the texts as a whole, the results from testing the lowest passing model texts against EVP were revealing (see 4.2). The difference in terms of richness of language and high-proficiency expressions were, in
lack of a more suitable adjective, staggering. Granted, this study, did not analyze “expressions” in the sense of phrases or clauses but the words in their context can tell a lot about the proficiency level. An example is the word “Leave” (in N2) that I analyzed as a B2 word using the EVP as a tool. The word “leave” as in to leave the room is naturally not a very advanced word, in fact it is A1. However, in the context of the Norwegian model text it means to leave something behind after death which makes for a much more impressive expression, equal to a C2 proficiency. Evidently, the range of interpretation is very big, as it tends to be with language testing.

And although there are words on the list such as priest and himself that were misspelled (prest and him selv) and there is an argument to be made for not including them, the difference between N2 and SE- is substantial. The Swedish model text is very underperforming in comparison and shows a very poor vocabulary. To give a reference, the in-document-commentary on SE- says that “[w]hen it comes to the breadth and capacity of the language the pupil shows an ability to express himself simply and with a certain degree of vocabulary (hobby; almost; weekend; fail; subjects). (Swe Guidelines 2014) Without putting any value into either case, one can conclude that EVP and by extension CEFR does not find that list of words noteworthy, only hobby and fail make the cut and they are both considered A2, which is remarkably below expected level of B1 for 9th graders. This, along with all the previous observations, would strongly move to indicate that the Swedish national test is not up to its own set CEFR standard.

In an effort to see the bigger picture of this and not just the lowest graded texts (although they inevitably take the spotlight), Norwegian model texts seem to follow the same tendency of focus on content as their initial guidelines suggested. Noteworthy in Table 6 is the rise in wordcount parallel to the rise in grade of the Norwegian model texts, suggesting that longer texts get better grades. In contrast, the Swedish model texts do not share this tendency, which is in line with the language-focused approach mentioned in Method and Material.

The results of N4 give some relevant insight. The text is 549 words. The remaining of assignment 3B is considered C2. However, with only first half accounted for it is deemed C1 and with only the second half it is B2. This implies that W&I takes a holistic approach. Although the examples in which removing some specific word or adding some specific word changed the over-all grade speaks to an analytical approach (relate to the EVP hierarchy of words) Sources in N5 constitute words that show a spike of knowledge affecting the over-all grade. Although a teacher would not normally
consider the level of language in an article that the pupil was given and simply copied its name into the end of his or her text.

N3 was split in half and showed very different results in Write & Improve. What does that mean for the validity of Write & Improve? Does it imply an increasingly holistic focus? Without insight into the evaluative coding of the software we cannot say definitely. Perhaps by looking more closely at N2 and SE- and their vocabulary we can be given some answers. But first, let us account for Table 6 and its rather shocking overview.

The initial results of the Write & Improve assessment of the texts show that the Norwegian model texts hold a significantly higher CEFR standard than the Swedish ones do and that, in fact, the Swedish ones are below European standard. The results, as they stand without analysis, would be devastating to the Swedish national test developers. And so, the premature conclusions to be drawn are 1. Norwegian model texts show higher proficiency than the Swedish ones across the board, and 2. The Swedish model text for lowest pass (SE-) is an A2 text, which falls short of Sweden’s own set standard B1. What does it say about the two countries when Norway’s Write & Improve results uphold the Swedish CEFR standards from Figure 1 but Sweden fails?

Before we address that question we must question the results and their accuracy.

5.3 Validity of Write & Improve and Other Considerations

Can Write & Improve be a valid instrument to discern English proficiency? Does it accurately measure “what is intended to measure” (Hughes 1989) and is it “valid, accurate or well represented” (Newton 2013) when it comes to CEFR?

All texts in consideration, there are some abnormalities that discredit the validity of Write & Improve. For example it does not recognize anonymizations and in the case of SD- and SB changing 3 or 4 anonymizations to something, like XX (place) into Paris improved the over-all grade by one proficiency level. The principle of the observation made for N2 should be applied to all texts, which was to view the model texts through an examiner’s eye and adjust to that end. When adjusting accordingly, SD- and SB are heightened one grade by exchanging the anonymizations with conceivable replacements, asforementioned and such a shift can be. In the case of N2, removing the quote had no effect, but removing the sources of N5, as mentioned in the observations, changed the grade from C2 to C1, meaning that the language shown in the titles of referenced articles that in no way could be accredited to the pupil, affected Write &
Improve’s analysis. The difference from C2 to C1 is more significant than one might think, if one considers the language used to describe those levels; a C2 written production is a “smoothly flowing, complex text” (englishprofile.org) (see Table 3) on a mastery level and a C1 written productions are “well-structured texts of complex subject”. A complex text is vastly different from a text of a complex subject if you look at it harshly, the former applying to all the language produced and the latter talking about subject and indicating the over-all language is not necessarily complex. Additionally the impact those mere 42 extra words had (in the sources of N5) begs the question of how influential the English Profile hierarchy of words is. Both the EVP and a focus on the analytical approach could explain the increase.

Influence and focus points are relevant when reviewing validity. We know that for Write & Improve to be valid it needs to “examine” in accordance to CEFR-levels but since it is obviously not a perfect software yet, where does its strengths and weaknesses lie? To understand its priorities and way of evaluating written text irrefutably also exposes strengths and weaknesses in the very results at display here. Is complex vocabulary more important than clarity or intelligibility, for example? Clarity is never an issue in the Swedish text; however, despite its richness, the Norwegian one is sometimes entirely unintelligible. Example: “he is a good teacher, man how to solve for the strict things out of lost the people” (N2, task 1). Pay attention to the word *strict* in this context and to the different definitions from EVP featured below.

- **strict** adjective PERSON 51
- **strict** adjective RULE 52
- **strict** adjective BEHAVIOUR 55
- **strict** adjective EXACT 57 (EVP)

This one is particularly interesting as it ranges from B1 to C2 and from the information given in the model text it actually is not possible to tell which function the adjective is supposed to have. It is a high-proficiency word but it does not mean much when occurring in such a poor context.

Write & Improve is not perfect but one can argue that the difference in quality between the Norwegian and Swedish model texts is significant to the degree that implies it may be true that the Swedish texts do not meet the standard – at least not the European standard. Although it must be reiterated that the Norwegian texts are longer which makes them able to “cover more ground” linguistically speaking – the length could also make the text prone to more mistakes. The model texts that are the most comparative or meaningful may be the ones representing the lowest pass – the ones
answering the question “what should a 15-year old know?” which according to the Swedish CEFR goals (Figure 1), of which Norway does not have an official equivalent yet, what a 15-year old should know equals the proficiency levels of B1. The “poorest” texts from both countries are rather similar in length yet show different result. Norway (N2) holds up to the Swedish set standard B1, while the Swedish text (SE-) fails at A2. Additionally, SE+ and SD- are both also found to be A2 texts, making things more conclusive.

Except for SE- and N2, however, most of the Norwegian and Swedish texts respectively do not match in length and for that reason, if one wishes to really see the impact length of text might have, a modification of the texts might be needed. This type of segment analysis could be interesting to see in the future, especially if by chance, softwares like Write & Improve are something examiners and educators will rely increasingly upon.

Testing the quality of these model texts is problematic and there are no previous studies that have made the comparison I have. The human element has been covered but in addition to that, such texts or tests are not as easily evaluated because they are not countable (although I have made a fair attempt at it) as a math test would be. As mentioned in 2.3 there is a lack of tasks or items that measure knowledge in free texts like these (Fulcher & Davidson 2007). This puts a big responsibility on the teachers that need to assess pupils knowledge and performance as shown in tests, and logically an even bigger responsibility on the constructors of the tests.

The validity of teachers has been questioned throughout most of this study and the validity of Write & Improve has also been questioned and in some regards disproven through simple tests like removing anonymizations, as I have.

The crucial question is how they compare to one another. This is interesting for the educational system as it stands today with nearly exclusively human examinators and the educational potentials of tomorrow with new technology.

For a long time now computers have dominated the field of reliability, which roughly means upholding consistency in results (Shillingsburg 2016). A simple example of this is how a calculator can consistently count many times faster than a human and without error. The calculator also upholds a high validity as it accurately represents our mathematical system. This can be compared to the unreliable grades the Swedish Schools Inspectorate noted in 2013 (see page 5-6).
Language, however, is another story and has more cultural components that a simple software does not pick up on. Write & Improve and its software developers work around this by using so called “training data” by which it constantly learns from users input. (writeandimprove [www]) Of course, if we are comparing, the same way a teacher or test creator learns by assessing texts and her/his performance becomes more and more valid. Let us call our human/teacher/test creator Bob to make things simpler.

Sundqvist would argue that Bob could be experienced, inexperienced, Swedish, Norwegian, professional, unprofessional, attentive, neglectful or simply hungry like Danziger’s judges (2011), and, is thus, prone to extenuating factors that affect his approach when assessing and his validity (Sundqvist 2017).

Write & Improve is not prone to such extenuating circumstances unless one were to consider malware. Bob rapidly improves his assessment and achieves greater reliability and validity within the first 2-3 years of work experience and then stagnates and almost entirely stops in development (Sundqvist 2017). Meanwhile, Write & Improve is, as its name suggests, constantly improving (writeandimprove [www]). This inevitable data input and storage should hypothetically make Write & Improve an immaculate examiner should one quantify times infinity but this study is primarily focusing on the present.

For Bob to really perform he would have to be “adept” as Lowe called it (1986) or “socialized” as Fulcher & Davidson phrased it (2007), into a defined environment or institution that sustains knowledge and aims to perfect it. Bob might not be adept and that would affect his validity. Write & Improve is its own environment, in a sense, and builds upon its foundation of data, effectively making it adept in Lowe’s definition of it all.

So far the comparison weighs in the favor of Write & Improve in that, despite its obvious quirks and flaws it may be just the tool to propel language testing and evaluative techniques into a new era.

However, language testing and validity is more than just consistency and memory. Lowe (1986) would also argue that to understand guidelines (model texts, for the purposes of this discussion) one must be able to see them as a framework and not an absolute truth, but something to relate to. When evaluating the quality of a text one can do so analytically and holistically and while a holistic approach may be more popular Erickson (2009) would not omit analytical as an option from this discussion. How would Write & Improve draw the line between analytical and holistic and which
impressions weigh more heavily for the final grade of a text? Would it be able to strictly follow a grammatical rule at the same time as considering the cultural component? For example, would the software know that “thru” is a serious misspelling in a formal text but acceptable for a more informal text in an American setting where this mutation has occurred and become somewhat common practice? It probably knows it as it is exposed to it in its training data, but it would not know when to apply such knowledge.

When it comes to sensitive nuances like this, Bob, although with a much lower minimum performance, also has the highest potential of finding a balance and such a balance would then generate a superior validity.

To summarize, while the strengths of Write & Improve very much lie in it not being human, so does its flaws. With its enormous database and quick assessments it may already be a preferred method to a poor-performing Bob but it is not yet a match for well-performing Bob.

5.4 Contextual consequences
The results as they stand without analysis, would suggest that a change needs to be made, suggestively increasing the three-fold role that Swedish teachers play in the process to a four-fold one, like Norway in an effort to consolidate with the teachers that, according to Sundqvist (2017:6), have a unique understanding and insight.

However, these results carry with them a lot of baggage and valid criticism. Results like these, while interesting as raw statistical indicators, need contextualizing. This last section of the analysis is meant to further contextualize how we can understand these results and observation and their impact.

The Norwegian pupils had beforehand read material they were supposed to reflect upon from which they can lend expressions and vocabulary. The Swedish pupils are also usually prepared before national tests but not afforded the luxury of knowing the subject they are going to write about. This gives Norwegian pupils an advantage that at least partly explain the substantial differences. The Norwegian model texts are, in general, more inclusive and transparent with their four-fold role and preparations, making both teachers (directly) and pupils (indirectly) a part of the process of standardized testing. This is not to say that being inclusive is necessarily a good thing. Lowe (1986) and Fulcher & Davidson (2007) both thought inclusion was the best path for improvement and Sundqvist (2017) actually bemoaned the lack of such an inclusion in the Swedish context, his argument being that nobody is better suited than foreign
language teachers with their unique understanding of the effect and implementation of the tests.

Sensor training will always be an extraneous element that this study cannot properly represent but that inherently might give Norway an advantage when it comes to adeptness. Becoming a part of a larger structure for this purpose ought to make the teachers more adept in Norway and being more adept also means being able to handle the inconsistencies this paper has observed more aptly.

The above arguments may all be true and reasonable in theory but this study must make a counterargument for the sake of reality. Inclusion into adept institutions may be a perfect architect of validity but that does not necessarily improve things if the blueprint is faulty to begin with. Broek & van den Ende (2013) and Erickson & Pakkula (2017) have both demonstrated how inconsequential and loose both Sweden and Norway’s relationship to the CEFR is.

It is possible that Norway’s educational system has a stronger communication and reach and that the model texts are better understood since teachers can also take part in the process. However, Ellingsund (2009) would argue that this could also be a negative thing.

Communicating something that is not intended or entirely correct can have a strong washback-effect and be particularly hard to counteract as it spreads through society in a negative cycle (Ellingsund 2009) where pupils of the established doctrine become teachers and then, respectively, teach his/her pupils the same thing. Broek & van den Ende (2013) concluded that implementation of CEFR is interpretative and this allows countries (like Sweden and Norway) to do so in half-measure.

This study makes the claim that, because of the wash-back effect and representation the model texts have and has had, implementing a new, more clear policy will be challenging and a long-winding struggle. We are in an educational era where standardized tests are increasingly seen as the measurement of communicative competence. Therefore, this study also emphasizes the importance that such changes in policy are integrated into the model texts (at least for Sweden and Norway).

The model texts are the subjects of this study because of what they represent and this study has tried to understand their role by examining contextual factors. But even if one were to disregard all contextual factors, the chosen model texts still matter since examiners assess texts with their help and language stands on its own, open for interpretation. The sensor training may to varying degree have an effect on how sensors
interpret the model texts but that is just the thing - it “may” and is not reliable; not as reliable as the fact that humans are affected by things they read, which is the aforementioned problem with teachers as examiners and the washback-effect. If they are staring at this very document when they are evaluating texts – that will undoubtedly have an effect of some significance. Consequently, it is important to consider.

The Swedish national test leaves the field open, leaving more room for mistakes or for creativity, depending on how one sees it, for the pupils’ own vocabulary. The Norwegian one sets the stage in another way and these things have likely been considered by the two entities when deciding the quality of the model texts. Do the Norwegian model pupils “write straightforward connected texts” as was requested in Table 3? Barely, and the initially impressive vocabulary is somewhat tainted since the pupils do not know how to use many of those words, making it less impressive but still acceptable. The Swedish texts on the other hand are simply too poor in quality, and even disregarding Write & Improve, the segments highlighted by the official guidelines as exemplary are actually consisting of A1-A2 words according to the EVP.

Norway barely lives up to CEFR standards and Sweden does not live up to them. However, even if both texts were at the same level of proficiency and directly comparable (which they are not), there might be a case for arguing that the results from the Swedish texts are more significant as they have officially claimed to be corresponding with CEFR, whereas Norway has taken the humble “we’ll get there”-approach. Cowardice, perhaps, but less blatant of a failure. In fact, the contextual considerations made in this paper inform us that the Swedish and Norwegian model texts compare rather poorly. They are not aiming to do the same thing and so the differences in quality of the texts are not relevant in the original rivaling sense of which text is the best, which pupils are the best. They are relevant because they inform of a big difference in approach to the national test and this paper observes and warns that changes may be warranted. There are many things that influence how a teacher examines a paper, their adeptness absolutely, their having or not having of lunch, apparently, but if one turns the argument on its head and talks about how the national test affects the teacher one can quickly see how great the washback-effect of two so different approaches, in a syllabus that increasingly values standardized testing as part of the final grade, can become.

Hypothetically, the Norwegian teachers become anxious that their weaker pupils will have to learn what an emancipation proclamation is or, even worse, feel forced to
teach their pupils how to take a complicated text full of difficult words, fill in the gaps and simplify it using their own words. Meanwhile the Swedish teachers comparatively and jealously glance at their Norwegian colleagues, concluding that the Swedish model texts are far too simplified and the Swedish days at the top of the list of non-native English speakers are counted. Additionally, and this I can speak to as an observed truth having talked to many English teachers in southern Sweden, they might feel that they despite established practice ought not show the model texts for the grade E to their pupils, in fear that the low quality will demotivate them from improving their English at all.

The initial findings of Write & Improve naturally sparks interest and debate and holds up when you start to look at the richness of the texts in more detail as I have done. It does to some extent calculate the quality of a text. Its results are useful for comparing texts because one can count the qualities quantitatively rather than trying to compare the texts relatively and rather subjectively to grade criteria. Nevertheless, it is a blunt calculation, blind to many considerations such as quotes, prerequisites and instruction making it too unreliable at times. To quote myself from before: “[t]hey[Norwegian model texts] exhibit bigger vocabulary but how much of that is due to pure quantity and how much is due to the nature of the exam?” – these kinds of relative questions partly put Write & Improve and likely many other similar softwares to shame.

The claim has been made that the Swedish model text (SE-) is an A2 text, based on the holistic Write & Improve calculation, the EVP- analysis and the observations I have made in connection to those. This is to say that the text is below the expected level of proficiency, but how important that is and what consequences it may have is yet to be answered.

Ellingsund (2009: 2, 79) would argue these strong tendencies in a model text creates a washback-effect, as it sets a standard by which the whole system could be affected. “Exaggeration!” someone gasps. “It is but one little text, it is not going to affect how a whole system operates” another, slightly more rhetorical person exclaims. The importance of the observations of this degree project should not, however, be underestimated, namely for two reasons – the lack of validity and the impact of an official standard. As Swedish teachers come from different places, different schools of thought and are not bound to any strict method of assessment, they are also not adept in the way Lowe defines the term. I can tell from my own experience as a full-time English teacher that the validity of the grades to the somewhat loosely defined common
standard is entirely up to the individual teacher and varies greatly depending on each teacher’s personal interpretation and skill. And so, the growing importance of the national test, a mistrust-driven policy as it may be, become the guiding lighthouse, the unquestionable truth, as to what level of proficiency is to be required for a pass in the 9th grade. SE- represents that “truth” and, as this study has discovered, has some serious flaws. This to mean that when SE- exhibits 80 % problematic sentences and a lackluster vocabulary in a very short text, it creates a washback-effect. Teachers in Sweden then adopt those ideas since they come with such authority and relate to those model texts whenever they are setting a grade, consequently lowering the standard of what is considered acceptable English in Sweden and, perhaps, affecting Swedish competitiveness on a larger scale.

6 Conclusion

Language testing is a complicated art, and experts seem to grasp for words to aptly summarize it but maybe it is best summarized as “complicated” with the many frameworks, criteria, human elements and other contextual issues to account for. Despite these complications, this paper has reached and boiled down a set of more-or-less accurate answers to the research questions. This study aimed to study the model texts for the national standardized test in Sweden and Norway, their comparability in terms of context and language. The idea of the study arose with new technology for determining the quality of a text, Write & Improve, which also served as the main way to conduct the language portion of the comparison between the two countries. A detailed analysis of the model texts, using Write & Improve, the EVP, personal communication with stakeholders, national contexts and statistics has been conducted and garnered result. These results have been given further context through theories on language testing and a deeper understanding has been reached, on how standardized testing in different countries can shape, be shaped, incorporate and desillusionate our educational system.

The results of Write & Improve, what they look like and by which notions they seem to be produced, has been a core part of this study and has been discussed exhaustively from different angles. It upholds reliability but struggles with validity, as was exemplified by the inconsistencies found in the results for the different model texts (see section 5.1) and the theoretical comparison of validity between the software and humans.
(see section 5.3). When it comes to evaluation, Write & Improve is, at least for now, best as an overview or exploratory tool – something to start off with and then take further with a deeper manual analysis. That said, with technological progress the established doctrine is often challenged; this study has been a small example of that, as Write & Improve, although faulty, was the factor that exposed a weakness in the traditional language testing that surprised test creators and sparked a debate. These softwares, if Write & Improve is any reference point (which, given its status within CEFR and level of sophistication, it should be), are perhaps not mature as of yet to become the checks and balances on test creation. However, it may only be a matter of time before they become invaluable to modern education.

The Norwegian model texts are not as good as initially thought, the preparatory vocabulary being especially deceiving and this paper is uncertain whether or not it holds up to the (Swedish) standard of B1. The Swedish model texts, more specifically the SE-, are not up to standard. While the observation made by Bruun & Diaz using Write & Improve falls short in the sense that the software is not reliable enough, the indication proved accurate and the Swedish model texts lack in validity. SE-, the model text that sparked the initial interest for this degree project, proved to be as poor as it appeared. Despite SNAE’s outspoken focus on comprehensiveness, and the text is comprehensible, it is far too poor to hold up to any other standards, far too simple to be representative of one of the best non-native English speaking countries of the world and certainly not up to the B1 standard that has been set for Swedish 9th graders.

All these points are not meant as a mere slap on the wrist for the test creators. They need to be seen in the context and the gravity of it all. With the sensitive nature of language testing and validity and the subsequent washback-effects, discrepancies like the ones observed in this degree project are dire.

Two, so promisingly similar countries and national tests, proved very different indeed, leading one to ponder three things.

- The applicability of the saying “Looks can be deceiving”.
- What a study like this would look like at an even larger scale, suggestively with more European countries (the EU and Europe being a region that for many decades now has aimed for cohesion and transferrable systems and measures). For future research, one should compare the methods and levels of proficiency that each country exhibits in their respective version of the national test and explore the possibility of a conjoint effort of cohesion around the work that
CEFR has already laid out, and how that might look in practice. As we learned from Erickson & Pakkula (2017), putting ideas into practice seems to be the greatest challenge for the CEFR.

- It will be increasingly important to standardize, digitalize and update language testing procedures to stay internationally comprehensive and competitive in a world where English remains the lingua franca.
Reference list

Primary sources

English Profile, Vocabulary [online]

http://vocabulary.englishprofile.org/staticfiles/about.html [Accessed 12 December 2017]


Guidelines Sweden – Skolverket (2013): Ämnesprov, läsår 2012/2013:
Bedömningsanvisningar, Part C, årskurs 9. Focus: Writing

Udir.no (NDET) 2014 ”Vurderte eksamenssvar i engelsk etter 10. trinn ENG0012” [online] https://www.skolverket.se/skolformer/forskoleklass accessed 2017-11-29

Write & Improve by Cambridge University [online] www.writeandimprove.com
[Accessed throughout the period of November 2017 – March 2018]

Secondary sources


Burman, M. Interview (E-mail exchange 16/10 – 23/11 with Matilda Burman, test assessment manager at the The Norwegian Directorate for Education and Training)


https://rm.coe.int/168045a0cf

Council of Europe, CEFR: Structured Overview of Scales [online] https://rm.coe.int/168045b15e [Accessed 29 October 2017] (No given year)


EF (Europeiska Ferieskolan) [online] https://www.ef.se/epi/regions/europe/sweden/ [Accessed 6 January 2018]


Erickson & Pakula ”Den gemensamma europeiska referensramen för språk:”
Lärande, undervisning, bedömning – ett nordiskt perspektiv” Acta Didactica Norge
DOI:  http://dx.doi.org/10.5617/adno.4789

Fulcher, G & Davidson, F (2007) Language Testing and Assessment


Hellekjær, O-G. 2011 ”Videregående opplæring: Erfaringer med eksamen i engelsk” Bedre Skole Nr. 2


Newton, P. Lecture (2013) Oxford

NDET (udir) ”eksamen SPR3008 Internasjonal engelsk”.


Personal communication, Bruun, S. Phone conversation took place in October 2017.

Shillingsburg 2016 “Understanding Validity and Reliability in Classroom, School-Wide, or District-Wide Assessments to be used in Teacher/Principal Evaluations”

[Accessed 19 October 2017]

SNAE/Skolverket, förskola (2011) [online]


Skolverket pre-school
classes[online]https://www.skolverket.se/skolformer/forskoleklass [Accessed 2017-11-29]


Swedish Schools Inspectorate (Skolinspektionen) 2013: "Olikheterna är för stora: Omrättning av nationella prov i grundskolan och gymnasieskolan"

Udir.no [online] https://www.udir.no/eksamen-og-prover/eksamen/administrere-eksamen/ [Accessed 17 November 2017]

Udir.no https://www.udir.no/eksamen-og-prover/eksamen/administrere-eksamen/

Udir.no (NDET) (2017): “Eksamensrapport - tilbakemelding fra sentralt gitt skriftlig”

Udir.no/fagfordyelsen https://www.udir.no/laring-og-trivsel/lareplanverket/fagfordyelsen/ [Accessed 17 November 2017]
Malala Yousafzai is 16 years old, a Pakistani girl. She was a teacher in Pakistan and her father was a teacher, so the Taliban asked her to stop teaching women, but she thought women have powerful roles in society. The Taliban shot her in the head, but she continued her struggle in the world and wrote a book called *I Am Malala* about girls in education.

Malala had become a famous girl in the world by the equality of men and women, and Malala received Europe's biggest human rights prize in 2013. Malala, who is one of the major prizes, received many other international personalities such as South Africa's Nelson Mandela and Aung San Suu Kyi from Burma. While Malala is the first Pakistani who got this prize, she is the youngest girl who got this prize.

Malala took the decision and stood up once for the equality of education among men and women. The women have more power to give their life for all world's women and she writes the book *I Am Malala* and spread it all over the world.

Because the struggle person they got both sides of positive and negative. At the negative side most of them got imprisoned or killed and died, but at a positive side they leave a big history to the world and they their time and their life and save most of others.

Abraham Lincoln was the 16th president of the United States of America from 1860. He was 43 years old and the youngest president in US's history. He answered questions also in that divided population and he helped the black people and protected them. He issued the memorable Emancipation Proclamation in 1863. Abraham Lincoln was needed to see the equality of people among the blacks and whites.

He said, "Whenever I hear anyone arguing for slavery, I feel a strong impulse to see it tried on him personally."

Abraham Lincoln was in Dallas, Texas, and he was driven with his wife in an open car at the people's way and they were shot and killed. Abraham took the decision to protect the slavery's life and he went to give his life to die and to take the risk himself.

Nelson Mandela is a role model for many people in the world, because he fought against racism. Racism was a big trouble and there was a lot of differential treatment because of the people's skin color. Nelson Mandela, who was a black man himself, who wanted to stop discrimination against black people. After being imprisoned for 27 years, he slipped easily into the role of president. Mandela is especially a big role model for black people, because he made a big difference and he is a big part of what made racism and discrimination stop.

That's what makes him an important person for a lot of people in this world. A person I look to as a similar role model is Emmeline Pankhurst. She was a strong British suffragette who fought for women's vote and for female civil rights. She is a role model for me, and every other woman, or at least she should be. If it wasn't for her, it isn't sure that we have been treated the same and there isn't sure that there is has been equality between the peoples gender. Emmeline was a fighter against sexual discrimination, and did a big difference for women in the world. That's why she is a role model to me.

Dr. Martin Luther King Jr. was a civil rights leader famous for his *I Have a Dream* speech, given outside the Lincoln Memorial in Washington, D.C. in 1963. The main message in his speech is freedom. He meant that skin color and religion shouldn't matter. I agree with him, because nothing of that should matter at all. People are different, and everyone should accept that instead of abusing people because they are themselves.

In his speech he uses good words, and he talked high and clear and that made him a good speaker. He used words who describes his meanings very well. He used many worlds that I don't really understand, but that's because I'm Norwegian. So
for thousands of people who listened to, it was a very good speech with good words. Luther talked high and clear, with a lot of (innlevelse), who also made his speech very good.

362 words. Assessed as C1

Task 3D

Moving day

The painting "Moving day" from 1967 seems to be about a black family who is moving to another country. When a person who is different moves to another place, it’s harder to become a part of the people how is already living there. That’s because they don’t know anybody at all. In this painting the black kids have a cat, and the white kids have a dog. I think that is a detail with a lot of meaning. The cat and the dog are different, and they sometimes don’t like each other because of that. Dogs and cats can be friends too, but sometimes the dog can be (avstøtende), and not accept the cats for who they are. That is exactly the same as black and white people do to each other.

The little girl who is moving in to the new house is scared. I think that’s because the white kids is boys, and they are older. Another reason why she is scared is because the white boys don’t look friendly at all. They look at them like they are some animals, because they have dark skin. They are kids, so they don’t know better, and that’s why they look down on the new kids. The little girl is thinking that they aren’t good enough for the boys to be friends with them. They are black, and they aren’t good enough because of that. She thinks it’s sad, but she can’t do anything about it at all. They are different, and it’s nothing to do to change that.

The texts are about civil and human rights. This painting is about that too. The reason for that is skin color. The painting is about equality between black and white people. I think skin color shouldn’t mean anything at all. People are people. Race issues have been a big problem, especially in the USA. This is better now, but there’s still a lot of racism in this world. I think that’s something we should co-operate to get rid of.

N3ii

Task 1

He was a calm and nice man. He codes have brutal political discussions but even after a hard political debate, he would give the other politician a unifying handshake. He fights the battle for black rights and even after 27 years in prison he was ready to go for the goal.

Most people would come out and be hateful and want to take revenge on society after 27 years in prison, but not Mandela.

When he came out he were filled whit courtly, humble and good-humoured. And all this courage and the way he always kept on fighting is why he is a role model for so many people in this world! Today I don’t think there are any one we could compare Mandela with. He was a hero for all the black people in Africa.

Today there are a lot of everyday hero’s, who helps people that need someone. The human right isn’t getting taken good care of. There is war, sickness and hunger and all because of bracing the human rights. Maybe there will come someone like Mandela and stop this pain that has grown in our world.

Task 2

The speech is about making black people have the same rights as white people, and the message I get is that we don’t need to love each other but let’s live together in peace, respect and with equal rights.

A message that we could need today also!

He talks in a more normal man kind of way.

Not in a high-class politician that thinks he is the fucking king of the world kind of way.

He was a man for the people and of the people, not for himself and a bag of money.

And I think the people felt that, and understood him as the man he was.

And all of that made the message clearer for them to see.

Task 3C

Difficult life situation

In the preparation text we read about Deo. He is a boy from Zimbabwe, he’s family wasn’t poor but didn’t have much money.
He is a young boy and had to escape with his brother. They had to escape because of soldiers killed the family and friends in the Gutu district where they lived at first. Deo survives the first steps on the trip because he hides money that his mom hid away before she got killed in his football. He found some over 150 million Zimbabwe dollars. And since the economy in Zimbabwe is so bad that isn’t much.

They go to a friend of the family, Captain Washington, he was a police officer. Deo wanted to leave his brother there, but suddenly the soldiers show up at Washington’s door step and the two brothers must flee to South-Africa. The trip is long, hard and dangerous. They even had to hide in a truck, to get to the border. They arrive in Beinbridge later that evening. Then they start to play football. People start watching them, and then some South African kids joined him. They picked times and start playing, no one of them knew football as the back of their hand, but the ball starts rolling. Deo forgot about all the pains of life. It was just him and the game. He found a plays of mind when he played that just felt good.

Romeo and Juliet Romeo is born in the family Montague, and falls in love with Juliet who is a Capulet. This to families hate each other very much. This story takes place in the 1800 century, back then family pride and wealth were very important. They’re love isn’t legal. So they marry in secret, and sneaks around like two little love birds. But one day someone finds out, and both family’s acts in anger. Romeo and Juliet knew what to do. Romeo flees into the night, when Juliet is in her room thinking. The two lovers dreed to sun meet again, then to take there life’s white poison. The meet at night when most people sleep. They had dreed to now die together. They drank at the same time, and were now ready to die. But as the poison killed Juliet, Romeo survived. He quickly started to cry but knew what to do. He pulled out his knife, and stabbed himself in the heart. He now had joined his life’s love in dead. Both these stories about Deo and Romeo have a connection that they have to flee from there normal lives to get what they want. For Deo it was freedom, and for Romeo and Juliet it was each other. And another thing is that there actions leaded them to find an escape from the pain in their lives. For Deo it was football and for Romeo and Juliet it was depressing but cute enough, death. And both their actions led to something good.

After Romeo and Juliet were dead, the two families lived in peace and harmony. They could now love each other and befriends. For Deo, he found a new start on life in a new country. He could make a better start for him and his brother. They also made friend when they played football that could help them and show them the ways of South Africa. Both story’s actions ended up being for someone’s better, even though it both were bumpy rides, and that’s the similarity’s in these story’s.

N4

Task 1 Short answer

Nelson Mandela were a role model for many people to the day he died, and after his death he has turned in to an even greater one. Mandela fought to the very last part of his life, and he never gave up on what he thought was right. He was beloved, and he did not care of his own suffering. He was, and he still is, a real hero. Mandela was a lawyer, and a leader of freedom. He was a guarding star to many people because of his courage. Nelson Mandela believed in peace, he inspired people to believe there was a light in the end of the tunnel, he made them believe in. That’s something most people find the hardest to loose, their faith in whatever they believe. You can take their home, their family, their food and even their human rights, but without the faith, most lives have a lack in value that nothing else can fill. Your faith keeps your head up; it makes you want to be something, not just another piece of the great puzzle we are living in. It makes you want to be something, for someone. Mandela made many people believe in themselves, and that’s something that may make you have the faith in surviving. He didn’t only inspire the black ones; he inspired the whole world to believe that everything is possible with a positive state of mind.

I know a person who admires him, she is a former prime minister of Norway, and her name is Gro Harlem Brundtland. Brundtland was born in Bærum in 1939. She educated a doctor, and also went to Harvard, so she had everything she needed to have a perfect life. She was the first female prime minister of Norway. I believe many people put their trust in her ability, but I’m sure she got a lot of haters as well. Of course people prejudices, she was the first female prime minister, but I think she
Imagine yourself being at the hairdressers shop, and you go out from there, not all happy about your new hairstyle. The next day at work or school you see this person staring at you. You just feel more and more bad about your hair because you know the person can see the difference from before. This is what I want you to think about. You are not bothering the other person if you compliment him or her. Just as you are walking around thinking about his or her beautiful hair, he or she might feel ugly about it. So just tell them if they look beautiful, they deserve it and want it just the same way as you. I’ve never received a compliment after getting a new hairstyle, nobody likes me.

Another thing you can do to make other people happy is to earn their trust. If they trust you, they know they have someone to rely on if something is wrong. Both you and I know how bad we feel if we have got no one to share our inner thoughts with if we actually want to. We know that releasing feeling of taking some of our problems off our shoulders. Both you and the person next to you desire it, remember that. You can make a change, just like he made a change in my life, I feel worthy.

For the same reason as complimenting, you should smile! It can make that one grumpy person have a change. Just as you are walking around thinking about his or her beautiful hair, he or she might feel ugly about it. So just tell them if they look beautiful, they deserve it and want it just the same way as you. I’ve never received a compliment after getting a new hairstyle, nobody likes me.

He Made my Day

He made my day, he smiled at me and that made me feel like I was in heaven. It’s weird to see how his little smile can be so meaningless to him, but make such a big deal to me. His eyes turn Chinese looking when he smiles, but they do still blink and shine. They are perfectly green with an inner ring of blue. They start blinking every time I look at him. Once he actually said hello, and I was happy in a week afterwards. My heart stops every time I look at him, and it makes me embarrassed.

Once he also held the door for me, and then I felt all these butterflies going crazy in my stomach. After that, he always gives me butterflies, but he doesn’t know that, I’m the girl in the wheelchair. I’m the handicapped girl with braces and only one leg. I’m that girl with no friends. My brain works all fine, but my body is ruining me. Actually, I feel like I’m separated in two, that I’m two half’s. However, he connects me, he makes me feel confident about being different. He takes his time smiling at me, without that fake laugh that often comes with when others do it.

Today is my first day of school after summer, and I haven’t seen him for months. I miss his warm smile and blinks. We are the same age, and today is the day, I’m in his class. We are in the same class, and I cannot believe it.

He drives me crazy, and I feel this silent storm inside me, just waiting for someone to let it out. When I get to know him, I’m right where I need to be, so let’s just say welcome to the future. Nevertheless, I know that I am hard to love; I keep most feelings to myself. I am stubborn as a mule, but I’m worth it. Now I can see the light in the end of the dark tunnel that represents my life so far.

Already now, before I’ve even spoke to him, it feels like I’ve known him for years. I trust him, and I can’t wait for us to be friends. Now is the time, the time for living. I’ve got a long way to go, but whatever doesn’t kill me only makes me stronger.

Imagine yourself being at the hairdressers shop, and you go out from there, not all happy about your new hairstyle. The next day at work or school you see this person staring at you. You just feel more and more bad about your hair because you know the person can see the difference from before. This is what I want you to think about. You are not bothering the other person if you compliment him or her. Just as you are walking around thinking about his or her beautiful hair, he or she might feel ugly about it. So just tell them if they look beautiful, they deserve it and want it just the same way as you. I’ve never received a compliment after getting a new hairstyle, nobody likes me.

Another thing you can do to make other people happy is to earn their trust. If they trust you, they know they have someone to rely on if something is wrong. Both you and I know how bad we feel if we have got no one to share our inner thoughts with if we actually want to. We know that releasing feeling of taking some of our problems off our shoulders. Both you and the person next to you desire it, remember that. You can make a change, just like he made a change in my life, I feel worthy.

For the same reason as complimenting, you should smile! It can make that one grumpy person have a way better day. Just a little thing as a smile can make you realize that not all that glimmers is gold. You can see it in the eyes of that person which is having a horrible day that you just made her or his day to something better, and that should give you some food for thought. In the long run, you will earn on this, because people will like you better.

We are going in to the classroom at this moment, and I just can’t wait. All these thoughts are running through my mind. I’m heading my seat, and the teacher places him by my side. I feel like my heart stops, and I can feel his body moving behind me.
“Well, isn’t it better if she gets a desk by herself, I mean, won’t she need more space for that chair?” he asks the teacher loudly. My whole world falls down, I guess it was just an act he played.

Task 1
Nelson Mandela has through his fight for equal rights, become a role model for many people in the world. Mandela fought against the apartheid system in South Africa. It was a system created by the white immigrants to keep the racial groups white, black and colored separate, and to give white people more advantages. Nelson Mandela thought apartheid was wrong, and meant that everyone should have equal rights. Mandela was imprisoned for many years because he fought for what he meant was right. Throughout his struggle for equal rights Mandela has for many people become a symbol of opposition, and a great role model.

I think a similar role model to Nelson Mandela could be Rosa Parks. She is famous for the bus boycott in Alabama, during apartheid in the US. She stood up against apartheid when she refused to give her seat to a white man when the white part of the bus was full. The rules said she had to, but she refused to do so. This action caused a lot of drama, and Rosa Parks was punished for her actions. Just as Mandela did, Rosa Parks protested against apartheid. And she has also, just as he, become a great role model to many people.

Task 2
In Dr. Martin Luther King Jr.’s speech, the main message is that everyone should have equal rights. All humans are born equal. People should be judged by the content of their character, not by the color of their skin. He wishes that one day, slavery will be abolished and everyone will be able to live together in peace.

The language Dr. Martin Luther King Jr. uses strengthens his message in several ways. First of all Dr. Martin Luther King Jr. starts every new paragraph in his speech with “I have a dream”. He could have said “I hope that…”, but phrasing it this way makes it sound much more powerful and strong. He describes freedom and equal rights as good things like “an oasis of freedom and justice”, and slavery and discrimination as bad things like “a desert state, sweltering with the heat of injustice and oppression”. He talks about how much better everything will be when his dream becomes reality, which makes more people want it to happen.

Task 3b.
Article.
Small steps to change
Not everyone can have as much impact on the world as they might wish to. Not everyone can be Nelson Mandela or Dr. Martin Luther King Jr. Not everyone can make a huge difference in this world, or change it all for the better. But everyone can make a little difference.

What do you do when you find out you can’t make the change you wanted to? Or when no one is listening to what you have to say about what you believe in? Or when someone argues against your dreams and beliefs?
People often have dreams about this; making change. They want to make a difference, and they want to matter. They want to change the world into a better place. Many people have dreams about stopping all war in the world, or making all racism end, and etc. But in a lot of cases, these dreams never become anything more than dreams.

After a while you might see that you won’t be able to make the change you wanted to. You can’t stop all racism. You can’t save all the lives you wanted to save. So you give up on your dream. Or maybe you forget about it. Other things matter more. Maybe you have a family you need to take care of. Your dream gets put aside. You don’t have time for it anymore. So instead of doing the things you wanted to do to help others, you start thinking “poor them, but there is nothing I can do.”

Or maybe you do try. You try standing up for your dream and for your beliefs. But there is still no guarantee that people will hear what you have to say. Maybe they don’t listen to you, and maybe they don’t want to listen to you. And after a while it starts feeling pointless. And as I just said, you end up thinking “poor them, but there is nothing I can do.”

Don’t do that. Don’t give up on your dream. There is always something you can do. Always a difference you can make. But don’t expect to change the whole world. If you can make a difference to someone, or change one thing for the better, then that’s great. But you can’t help everyone. And you can’t change the world over night. It’s all about taking small steps, one at a time.

Look around you. Start with the people closest to you, your family, friends, neighbors and etc. What can you do to make things better for them? Small things you don’t think matter often make much more difference than you think. Be nice and be helpful.
And if you were of those who tried, but didn’t get heard, there is always something you can do to change that. Find other people who share your dreams and beliefs. It is much easier to get heard and make a difference if many people stand together. There are organizations for nearly anything, which you can join. These will make it easier for you to participate in the things and cases you want to participate in.

And then the last question: what do you do when someone argues against your dreams and beliefs? Or if someone does something that is against what you stand for? Don’t get furious. Instead try to be smart about the situation. Explain your opinion to them or tell them why they can’t do what they’re doing, in a polite way.

For example: say someone is saying or doing something racist to someone because of that someone’s skin color or sexuality. Even though you might get angry, don’t start yelling at them. If you do, I can almost guarantee you that they won’t listen. Let’s say that the person making racist comments is a white man from Norway. Ask him “Sir, did you choose to be born as a straight, white man in Norway?” The answer is obviously no. Then you can ask “So do you think that this person chose to be born gay, or in a black family? No, he didn’t. And you treating him bad because he’s gay won’t make him any less gay. In the same way as saying racist comments about his skin color, won’t change it. It’s not a choice he’s made, and he shouldn’t be judged for it. He has the same rights as everyone else.”.

Maybe your dream was to stop all racism all over the world. You won’t be able to do that. But if you do what you can, if you try to help someone that is ever in a situation like this or a similar one, then you have made a difference. Maybe not to everyone but to that person. And small steps like this can end up making a big change in the end.

Sources:
Nelson Mandela Inspired Millions To Bridge Divides Seek Peaceful Change, extract from preparation sheet.
Strong civil and human rights voices, extract from preparation sheet.
I have a dream, extract from Dr. Martin King Jr.’s “I have a dream” speech, August 1963.

N6
Task 1
Nelson Mandela
Why is he a role model for so many people in the world?
Nelson Mandela was imprisoned for 27 years, only because of his political opinions. The Apartheid regime in South Africa did not want the black majority in South Africa to have any power of some kind, because then they would lose their advantages. Nelson Mandela was not liked by the regime because he was a black man, with political opinions. Most people would’ve been bitter and filled with hatred after 27 long years in prison, but Nelson Mandela smiled and was just as good-natured as he used to. Those years had only made him grow stronger. He didn’t fight against the cruel and racist regime by power – but he fought by words. And if a person such as that isn’t a good role model, no one is.

Who do you look to as a similar role model and why?
Mahatma Ghandi is a similar role model. The reason of why I think that Mahatma Ghandi is a similar role model is that he fought against the minority in India, the British. When Mahatma Ghandi fought with words and boycotting around the time of the First World War, was India still a colony under the British Imperium. Just like Nelson Mandela in South Africa, he fought not by power – but by words. The British didn’t like him, just like the Apartheid regime didn’t like Nelson Mandela. Mahatma Ghandi was also imprisoned by the rulers, the ones with the greatest power – the minority. And just like Nelson Mandela, Mahatma Ghandi and the Indian people won against the minority and got their country back. Mahatma Ghandi did the same as Nelson Mandela, only before him.

Task 2
Dr Martin Luther King Jr.
What is Dr Martin Luther King Jr.’s main message in this speech?
The main message in this speech is that he has a dream, and that dream is that the United States of America one day will live out the true meaning of the Declaration of Independence, that all men are created equal. Which means that his dream is that every person, no matter sex, skin colour, political meanings, religion – everything, are created equal and therefore are to be treated equal. But he was mainly talking about racism- because he lived in a USA where black people were looked at as inferior and he himself was a black man.

How does the language he uses strengthen the message?
When he starts the most important sentences with the same words, he uses repetition and that made it probably easier to catch up with what he was saying to the crowd. He used repetition as a type of means. With the “I have a dream…” before a great
sentence he can take a pause, so the crowds could think and wait in excitement for what that dream could be – and when he then continues with for example “… that this nation will rise up and live out the true meaning of its creed…” it’s great and powerful, almost as powerful as the strongest weapons – if not as powerful. These words are not forgotten, they’re a big part of the American history and I personally love the speech because of the great power in all the words he used.

Task 3B (fictional)

Prejudice

“Homosexuality is wrong!” My parents always brought me along on those demonstrations. They wanted me and my siblings to learn that we weren’t allowed to be gay. That we were supposed to find someone of the opposite gender, marry and then have children.

We learnt that we were not by any circumstances allowed to be in a relationship with someone of the same gender, or someone who didn’t want to marry. An important part of our childhood was to learn that everyone that weren’t good Christians as us, were living their life as sinners and were sentenced by God to live eternity in pain and sorrow.

Jimmy laughed and pressed me against the wall.

“So I’m a sinner then?” He smiled proud. “I guess…” I answered and looked deep in his warm, brown eyes. They always made me feel so good and safe. Every time I looked in his eyes it was just like falling in love with him again, over and over until I didn’t remember my own name. “Doesn’t that make you a sinner as well?” He kissed me on the neck and pressed my body closer to his. I nodded, slowly.

“Cornelius! Don’t look at them you might get their disease.” My parents always said so when they saw homosexual people in the street. But already then I knew it was too late, ever since the first time I saw into Jimmy’s eyes I knew it. I was born to be a sinner. My only purpose on the surface of earth was to be the faithful boyfriend of the according my parents – rebel. His curly, dark hair made him look even more like a Greek god than Zeus himself. Yes, totally, might get their disease. I’m blessed with this so-called disease that my parents are so very afraid of.

I remember it so good. How I had judged him the first time we met, how I had stared at him like he was weak and useless. Only because of what my parents taught me, that it was a disease and that it was wrong. But it didn’t take long before I discovered that I was just like him. I remember how painful the first weeks was, how I struggled to accept myself. Every time I saw Jimmy I would feel two things – intense love and intense hate. He could smile to me and try to calm me down, while I shouted at him and tried to hit him for making me gay.

And I cried a lot, cried so much that my eyes almost dried out. But after a while did I accept it, and learn about it instead, the prejudices put aside.

“But Cornelius, honey…” he started, gently stroking my back. I looked up at him, and felt that my face had gotten warm. He kissed me on my cheek. “What are you thinking about?”

My face got warmer and started to beat in the same speed as my heart. “You know…” my voice were trembling “… just before we became…” That absolutely astonishing pretty smile of his.

“Boyfriends?” he whispered flirty in my ear. I felt a chill down my spine. “Aye…” I whispered nervously back. “What about it?” he kissed my cheek again “I’m so sorry for everything I said and did and….” a tear suddenly trilled down from my right eye, and with his warm hand he gently stroked my cheek. “Oh, honey. It’s nothing to worry about, it doesn’t matter, and besides that belongs to the past.” Ashamed, I was so extremely ashamed of my behaviour back then. How could I’ve said such things to him, the kindest person on the surface of earth? “Cornelius, I can see you’re still dwelling on it.” I turned my face away from him, and I knew he was right. I was still dwelling on it, but I couldn’t help it. “Listen, I know you said some bad stuff, but you didn’t mean it. You’ve learnt from your birth at you are supposed to say things like that and mean it when it comes to homosexuality, and I don’t blame you, nor do I blame your parents – but the ones who wrote the bible – or started the homophobia, that person do I blame for making you say stuff like that.” He kissed me, intense and loving. “And I think it’s amazing how you got over your prejudices against homosexuality; you should teach other people how to do so…” “What do you mean teach people how…?” “Teach them how to take the same small steps as you did, because if everyone does that, the world would be a much better place.”

I guess that’s right, if everyone opens their eyes and stop judging people before knowing them, the world would indeed be better. Everybody should take small steps towards a great, common target – a world without any prejudices or injustice.

Swedish model texts

SE- Hey!
My name is NN, im’ 15 years old.
My hobby is to be with my family and my friends. I'm with my friends all of my time, every day after school, almost until the night.

Oh, I'm going to school every day, but not on Saturday’s and Sunday’s because then it is weekend and nobody goes to school then. In my school we have sixteen subjects and I don't fail in one subjects!

My life right now is wonderful, I got the most nicest mom and dad.

Trends come and go everybody knows that, but right now this technology we got today is extrem good!

Music changes a lot with the years. I mean when I was a kid I didn't know what screams and hard rock was, that was never on the radio. Today I'm listening a lot of music on radio!

One thing that is important to me is cars. I'm very interested by cars. I got three of them myself. I think that cars will be my interest all life, it's a big thing right now and the same in future.

Another thing that is important is school. School is important in that way because it's a ticket to your life and what you can do when you grow up and be a man and get a job and make money.

My life right now is very tough because we have too much homework and I do have to do it because we kids need to work at home.

When we need to work at home we don't get some time over to have fun like, look at tv, play floorball, be with friends.

I think we need to have more holidays so we can take it easy sometime.

When I have holidays I often play floorball and football because I like to be with my friends and football and floorball are team sports. Everybody in my team are my friends.

B) The technology has influence our world very much because the computer, the computer are inside our house and then we need to be inside to work/play with the computer.

If we are inside all days our health get bad and we will be sick because we need to be outside and breath fresh air to feel good.

C) One issue I will talk about is the politics. I think everyone in our world likes to work together to get our world more comfortable.

We need to destroy all cars that destroy our world. We need to get cars that drive on electricity.

We need to think on that people how comes after us, they want to have a wonderful world like ours.

Hi, my name is NN and I'm 15 years old. My favourite colour is blue, green and red. I like to play football, gymnastics, dancing, singing, play the piano and to be with friends. I have three brothers and one sister. My brothers name is A, B and C. My sister's name is D. I'm the youngest in my family. My mom's name is E, and my father's name is F. I live on X and in the city.

My sister D is married and has a son. She and her husband live on X, so from Monday to Friday I sleep at their house because it's closer to my school. And on the weekend I go home to my mom and dad in the city. The school that I'm going to is called Y. Sometimes I'm taking the bus to school or sometimes I'm taking the bike. When it comes to trends and how to dress or what kind of music you're supposed to listening to, I've always going my own race. I've always dressed how I wanted and listened to the music I've wanted to. I don't let anybody tells me how to be. For example in America, there is a big issue about the health. It's a lot of people who are over-weight. In America there is a lot of fast food and takeout. And many people are dying that kind of food. I'm thinking that they have to do it a lot healthier. And there is a problem about the nature. There are many people that are throwing trash in the nature. I think that people can take the bus, take the bike or walk to their works or where they're going to and not take their car or scooter. Because it's destroying our nature with all the smoke.

Hi, my name is Tom and I'm 15 years old. My favourite colour is blue, green and red. I like to play football, gymnastics, dancing, singing, play the piano and to be with friends. I have three brothers and one sister. My brothers name is Bill, Bob and Carl. My sister's name is Donna. I'm the youngest in my family. My mom's name is Edith, and my father's name is Felix. I live on the area and in the city. My sister Donna is married and has a son. She and her husband live on the area, so from Monday to Friday I sleep at their house because it's closer to my school. And on the weekend I go home to my mom and dad in the city. The school that I'm going to is called The School. Sometimes I'm taking the bus to school or sometimes I'm taking the bike. When it comes to trends and how to dress or what kind of music you're supposed to listening to, I've always going my own race. I've always dressed how I wanted and listened to the music I've wanted to. I don't let anybody tells me how to be. For example in America, there is a big issue about the health. It's a lot of people who are...
over-weight. In America ther is alot of fastfood and takeout. And many people are bying that kind of food. I’m thinking that they have to do it alot healthier. And there is a problem about the nature. There’s many people that are throwing trash in the nature. I think that peoples can take the bus, take the bike or walk to their works or where they’re going to and not take their car or scooter. Beacuse it’s destroying our nature with all the smoke.

Hi, my name is NN. I was born in 1996. I have dark hair and brown eyes. I live in X a small village in Y, Sweden. I also go in school here, Z. Before i get on horse riding but not so much anymore. Our horse got sick and we can still just walk with her. First she was sick about one and a half year, then she was okey for about maybe six month or something like that but now shes sick again. But know my intrestings are music and watch movies. To listen at music is great i think. I do it almost all the time. Listen on the radio or my phone. I have my own style and don’t really care about what other caring. I think it’s fun with cllhotes and stuff like that but i don’t need to have the latest things all the time. Sometimes for fun but not very ofent. Music i listen to what i like. I don’t think it is very important to listen to the latest songs. Technology, i have my own computer, a phone (Sony ericsson x101 i think it’s called, something like that. And my own TV, it’s about a year old or more but it works. You can only look at DVD films but that works for me. About my phone again i’m gonna save to an iphone 4. Most of the money im gonna pay myself and the rest a im gonna wish for Christmas.

One big probleme that we have is the global worming. I believe that earth will be destored one day. Not right now but in the future. We really need to do something before it’s to late. We also need to take care of the world and stop to go an like this. This probleme will also effect our health. I don’t know really if it’s gonna be a bigger probleme in the furture or if it’s going to be better. I hope it’s gonna be better, i really do.

Another probleme is the poor contries all over the world. It’s not many places who have it so good as we have here in Sweden. I mean if everybody in Sweden and after contries to just give 10 Swedish kr. (every family in Sweden or person). That would be alot of money and we could send this money to the poor contries and make sure that the get them. Right now i’m 15 years old and i go on a school named X in Y. Every week form Monday til Friday are always the same, week after week. I go up early in the morning when it’s dark outside, then i dress up, eats my breakfast, and then I’m off to school. About 6-7 hours in a day I’m sitting in my bench and are supposed to learn things, but i can tell you one thing, that after almost 10 years in school you don’t find it very cheerfull at all. You just don’t wanna go there anymore. But that’s just life i guess ...

After school anyway i always go home and just chilling for a couple of hours, to just have some little time on my own and can be able to think over the day and so on. And after this i always goes to the gym, beacuse to be able to train and see all of the results that I’ve seen is something i just love to do.

So now you’ve heard about my weeks when I’m in school, but i haven’t forgot the weekends of course. The weekends are my time to spend time with a person i care for very much, and I know that he cares for me to. He’s the best part of me. Different styles comes and goes over the years, and I’ve had many differents of them, so of course it has influed me. Know a days I can’t care as much as i used to, beacuse I’ve learn trough the years that i won’t let all differents of people have such big infact on me. I’m just gonna carry on and be me, cause this is me and I was born this way. Cancer is a big issue for me and many people all over the world. I think it’s just crazy how one thing can kill so many people and distroy lot’s of peoples lifes, forever. I’ve seen the damages on a person i care about, all the tears that he have left behind just beacuse he’s mother not here anymore... it’s just not fair. In the future I hope with all I have that they find a cure for cancer, the world would be such a better place without it.

C+

Hi future me!

When I’m writing this letter am I exactly 15 years and 27 days old. I’m going to my last year here on X. I will start on high school next fall and hopfully in a sowing class. I really love to sow and make my own clothes. It’s really fun to make them but it is also good for the environmen witch we have problems with these days. I have heard every since I was a kid that we humans don’t take care of the world the way we should do. We use cars that make our air full of bad things. It is almost like we just beilive that the next generation will take care of it. But there is just one problem. We are the next generation. And if we keep on treating our planet the way we do it will break down sooner or later.
I often think about how we can change the environment, and I sometimes believe that the positions could do more about it. But as it is right now can’t the even take care of the money in their own countries. Right now almost every country in Europe are having a economic crisis. I really hope that have been changed when you read this. And when you read this I really hope that you have got yourself some computer, and that you managed to get a good education. I also hope that you miss being me, because it rocks. I mean I have amazing friends, a fantastic family and really good teammates. When I say teammates I mean all the other girls in my floorball team. We are a awesome team spirit and that makes us into a very good team. We train three times a week and on the weekends we have matches. We are actually so good that we meet girls that are up to five years older than us and sometimes older than that.

If I should describe the world we live in right now in the 21st century with four words would the probably be fashion, money and technology.

I chose those words because the explains my life pretty well. Fashion, because I’m really into the whole fashion and trend thing. For me it’s really important to look good and dress well, and I think it is like that for many people. Money I chose because everything is about money. Everybody wants it and some people are almost willing to risk their own life just to get some money. And it’s also a big status thing. If your rich your worth more than someone how is poor. Technology is something that is everywhere in this century. People have flat TVs, iPhones, Macbooks, laptops, GPS’s and Ipods. It’s sick how we get all upsets with all these things. Just look at e for an example. I have Ipod in my ears all the time, my Iphone in my hand and my Macbook in my bag.

XOXO
A young you! <3

Hello! I am 16 years old and I live in a small city called X which is in Sweden. My parents are originally from Turkey but they moved here 30 years ago to get a brighter future for our family. I got a brother and a sister. They are both older than me, 19 and 17. My brother who’s the oldest one is having a break from the school for a year before he will decide if he wants to continue studying or find a job. My sister is going on high school and it’s called Y. I will will probably begin there after summer cause it’s simply a great school with nice teachers and people. I like to play video games both on the computer and my playstation. But one thing I love is soccer. I’ve been playing soccer since I was 8. I started to play soccer in a small team for a year until I changed to the biggest soccer club in our city which is called Z. I’ve been playing ther for about 6 years now and it’s going great. I also like to be with friend and go to the cinema once in a while.

Technology has influenced me a lot because we use it all the time. We keep in touch with friends and family in many ways like email, text messages, calls etc. We can work from home with just a computer. With technology I can entertain myself by playing a game or watch a movie on the computer, tv or on phone. Technology has also saved many lives because of the new machines they are using to see and solve problems. Without the technology we got at the moment life would be so much harder and it’s a big part of our life.

A big part of my life is my religion, Islam. If I didn’t have Islam in my life I would probably do bad things in my life. Islam has lead me in the right direction of life and it’s making me a good person and also want me to do great things for other persons. Islam is what making me a complete person and I can’t live without my religion.

A Example 9
My life is great. I go to a good school where I have many friends. From our classroom we have a stunning view over a lake. Sometimes I look out of the window, I see the lake and I fly away in my mind…

When I’m not in school, I study a lot and I play football. That’s an amazing sport – tough and fun! My dad do also play football so we use to practise together; discuss and watch the game on TV. Except football, I think Christmas and summer holidays are what I like the most. To discover other parts of the world seems fantastic to me.

On Sunday’s I’m a leader for kids and teaches them how to swim. It’s very fun, I feel like I’m their mother. They are so cute.

Right now I’m in school. What I like the most here is P.E., you know playing football, basketball, go swimming and such like that. The second thing on my list is actually English. I think it’s cool when you understand what people are talking about on TV and to be understood when you’re abroad.

I use to watch the programs “Top Model” and “Projekt Runway”. They’re a lot about
fashion and having the right look. I’ve got to admit that it has influenced me. I can’t tell you if it’s in a good or a bad way. I’m more like everybody else now, I don’t really got a personal style anymore. I’m more into trends now than I was before. But when it comes to music, I think I’m a bit different. Of course I can like new hits by popular artists in our time, but I also love hits from the 80:s and Bruce Springsteen etc. One of my favorite movies is actually an old one – Rocky. It inspires me to do whatever I want to and of course – train hard!

Today’s world is rough. Many people are poor, there’s a lot of violence and crimes, but one of the biggest problems is probably the question about the environment. The home of millions of people, animals and plants, are soon falling apart because of the people’s wrong decisions. Many persons knows about the problem but it’s just a few how really are trying to make a difference. But when the situation just is getting worse, everyone have to be a part of the projekt and try the hardest to get a solution. If not – the problem will grow like a virus and destroy our home in the future. In a hundred years, maybe there isn’t a planet called the earth…

Appendix B – Commentary on model texts from within the document of guidelines

N2 response
Content: Task 1 is to a certain degree in correspondence with the demands of the assignment and gives to a certain degree an answer to the question on why Mandela is a role model. The answer shows signs that different parts of the texts are haphazardly put together to fit the demands of the task. The text gives an unclear answer on the last part of the task.
Task 2: The text shows signs of being patchwork. The second part of the task is not addressed.
Task 3A: The text partly exhibits reading comprehension. The answer contains attempts at argumentation but could use quotation/reference to solidify an argument.
Text structure: The texts are to a certain degree adjusted to purpose, receiver and digital criteria. They contain simple patterns of text construction and are in part formulated with understandable structure and context. The answers contain some attempt at sentence cohesion like for example “and”, “so” and “but” and the texts are structured with a separate theme for each task. Task 3 has a main section without an introduction or ending.
Language:
Task 1: The text has a language that to a small degree communicates with the receiver. The answer has many words that are correctly written but the language is to a great degree lacking in conjugation and syntax.
Task 2: The language communicates to some degree with the help of patchwork.
Task 3A: The text communicates more clearly than the one in text 2 because the paragraphs contain some statements that are supported by examples such as in the last part of paragraph four.
Summary: The examination text show reading comprehension and that the tasks were understood. It (the text) has a simple content that to a certain degree matches the criteria of the task. It is to a certain degree adjusted to purpose, receiver and digital standardization criteria.
The answer shows a certain degree of correct spelling and punctuation that makes the text understandable. The answer shows “low” competence and is altogether deemed a grade 2 text.
(Translated by me from Norwegian)
The pupil presents himself and then tells about his everyday life in a simple manner (My hobby is to be with my family and my friends; I going to school every day). With the help of the task instructions a few examples are found in short episodes (Trends come and go; Music changes a lot). The pupil also shows the ability to choose his own examples (I’m very interested by cars; School is important) and briefly reflects on the future (when you grow up and be a man who got a job and make money). At some point he links between the texts parts (one thing that is important – Another thing that’s is important), which alleviates the reading experience. Apart from that, the content is not connected but rather appears as a streak of answers to questions and the text is completely lacking an ending. The content is fully understandable but relatively weak even if it at some points is developed with an explanation to the reader (I mean when; in that way it’s a ticket to your life). The pupil also to some degree appears aware of recipient and situation when he adds comments (it is weekend and nobody goes to school then; everybody knows that).

When it comes to the breadth and capacity of the language the pupil shows an ability to express himself simply and with a certain degree of vocabulary (hobby; almost; weekend; fail; subjects) as well as some functional phrases concerning personal relationships and concrete situations (come and go; it’s a big thing; make money) When it comes to comprehensiveness the fact that the use of verbs largely works contributes to it working in regards to tenses (when I was a kid; I think that cars will be; what you can do when you grow up) and congruence (nobody goes; Music changes; School is) despite some insecurities (I didn’t know; Too day I’m lisent). The pupil sometimes binds together groups of words with simple cohesive/linking words like but and and as well as, in some case, because. However, short clauses are most often linked simply using a comma in between (my life right now is wonderful, I got the most niceist mom and dad), which leads to flawed contextualization which disrupts the reader experience. The summarized evaluation is hitherto that the pupil shows an ability to briefly express a simple content that to some degree is adjusted to recipient and situation. The pupil can formulate himself in a simple, mostly understandable and relatively connected way. The text is therefore graded as acceptably fulfilling the knowledge requirements for the grade E.

(Translated by me from Swedish)