“Takes a little getting used to”

Perceptions of Finnish upper secondary school students on the digitalized Matriculation Examination for English

Sanna Tarvainen
Master’s Thesis
English Philology
Department of Modern Languages
University of Helsinki
October 2014
Table of contents

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

2 Theoretical background ...................................................................................... 5

2.1 Computer-assisted language learning (CALL) ................................................. 5

2.1.1 History and background for CALL .............................................................. 5

2.1.2 Why use CALL? .......................................................................................... 7

2.1.3 CALL methodology and devices ................................................................. 9

2.1.4 The use of ICT in Finnish upper secondary schools ......................... 11

2.2 Computer-based language testing (CBLT) .................................................... 13

2.2.1 History and background for CBLT .............................................................. 14

2.2.2 Why use CBLT? ......................................................................................... 16

2.2.3 Example WBLT realizations ...................................................................... 17

2.2.4 CBLT for the Finnish Matriculation Examination ............................. 19

2.2.5 Test usefulness qualities ............................................................................ 20

3 Research design .................................................................................................. 28

3.1 Research questions ......................................................................................... 28

3.2 Participants ...................................................................................................... 28

3.3 Data collection procedure ............................................................................. 29

3.4 Background for the example exercises ......................................................... 31

3.5 Questionnaire .................................................................................................. 32

3.6 Pilot study ......................................................................................................... 33

4 Results and discussion ....................................................................................... 34

4.1 Results about the digitalized Matriculation Examination ......................... 34

4.1.1 Attitudes towards the reform ..................................................................... 34

4.1.2 The learning environment ......................................................................... 40

4.1.3 Measuring and bringing out language skills ........................................... 45
4.1.4 The example exercises according to task type .......................... 52
4.1.5 Overall user experience...................................................... 65
4.1.6 Suggestions for improving the digital Matriculation Examination exercises............................................................. 65
4.2 Results about the relationship between foreign language education and the digitalized Matriculation Examination .................................. 67
   4.2.1 Results about current language education.............................. 67
   4.2.2 Results about required changes to English education.......... 69
4.3 Summary of results.................................................................... 73
5 Conclusion ..................................................................................... 77
References .......................................................................................... 80
Appendices .......................................................................................... 83
List of pictures

Picture 1 Students’ comments on example exercise 1: Multiple-choice reading comprehension ........................................53
Picture 2 Students’ comments on example exercise 2: Open questions .................................................................56
Picture 3 Students’ comments on example exercise 3: Grammar and vocabulary .....................................................58
Picture 4 Students’ comments on example exercise 4: Oral skills ......59

List of tables

Table 1 Data collection procedure ................................................................. 30
Table 2 Experiences on working efficiency in an electronic exam situation, according to gender ...................................................... 51

List of figures

Figure 1 The importance of the reform ....................................................... 35
Figure 2 Willingness to complete own Matriculation Examination electronically ........................................................................... 35
Figure 3 Willingness to complete own Matriculation Examination electronically, according to gender ......................................................... 36
Figure 4 Willingness to complete own Matriculation Examination electronically, according to own estimate of English skills ........... 37
Figure 5 The importance of testing oral skills in the English Matriculation Examination ................................................................. 37
Figure 6 Importance of testing oral skills in the Matriculation Examination, mean / Importance of digitalizing the Matriculation Examination ................................................................................. 38
Figure 7 Attitudes towards measuring oral skills in the English Matriculation Examination ................................................................. 39
Figure 8 Learning to use the software, according to own estimate of IT skills ...................................................................................... 41
Figure 9 Reliability of test software ............................................................ 41
Figure 10 Summary: familiarity or degree of enthusiasm towards the reform and test software, according to gender ........................................... 45
Figure 11 Versatility of testing language skills ........................................... 46
Figure 12 Authenticity of the example exercises ........................................ 46
Figure 13 Similarity between the computer-based Matriculation Examination and the paper-based one .................................47
Figure 14 Bringing out language skills, CBLT vs PBLT .................48
Figure 15 Learning while completing the exam ..............................49
Figure 16 Ability to concentrate, CBLT vs PBLT ..........................50
Figure 17 Experiences on working efficiency in an electronic exam situation, according to IT skills ...........................................51
Figure 18 Current English education ...........................................68
Figure 19 What students consider important in upper secondary English Education .........................................................68
Figure 20 Developing upper secondary school English education ....69
1 Introduction

Utilizing computers in language learning has long been a matter of discussion both in Finland and universally. While the 60s saw the discussion of whether the arising technology might have a place in the language classroom, the 21st century no longer asks whether to use these appliances but acknowledges to a growing extent the new responsibility of language teachers to educate themselves and their students in their usage (e.g. Blake, 2013; Chapelle, 2001, Walker and White, 2013). The surge in the use and span of social media provides numerous new opportunities for language use and learning, bringing foreign-language communication closer to the language student. Indeed, whereas authentic interaction in a foreign language used to require travelling to a foreign country, the closest opportunity for it is now to be reached for in the student’s own pocket.

The consequences of information and communication technology, or ICT, on language education are not limited to changes in language use. The developing ICT bears consequences to the teaching, assessment and evaluation of a foreign language, all of which can be made more up-to-date or efficient with the help of technological devices. What makes enhancing teaching and learning technology especially topical in the educational setting of Finnish upper secondary schools is the current development of the nation-wide Finnish Matriculation Examination towards an electronic format. This examination, the only form of standardized testing that is currently in place in Finland, is generally taken during the third or fourth year of upper secondary education. It is a high-stakes examination, as the grades have a profound effect on the students’ possibilities for further study. All students participating in the examinations participate in the exam for Finnish or Swedish as a native language, and, in addition, they choose a minimum of three subjects in which to complete the exam. Popular subjects include e.g. English, chemistry, psychology and mathematics (Ylioppilastutkintolautakunta 2014).

The advanced-level English exam, which is the topic of the current study, is the most popular subject of the Matriculation Examination when it comes to the number of test takers (more than 40,000 test-takers each year; see Ylioppilastutkintolautakunta, 2014). Despite offering insights to the development of the exam which touches the highest number of test-takers, this study will discuss and
identify issues which are relevant to the Matriculation Examination of foreign languages in general.

According to the current schedule of the digitalization project “Digabi”, initiated by the Matriculation Examination Board in 2013, the exams are to be made electronic starting from the year 2016. By spring 2019, the exams of all subjects will be in the electronic format. The aim of the project is clear-cut: to complete each of the five stages of the Matriculation Examination by utilizing ICT. These stages are 1) the preparation of the exams, 2) completing the exam, 3) the initial grading stage at the schools, 4) the final grading completed by the Matriculation Examination Board, and 5) the reporting, publishing and archiving of the results (Digabi, 2014). The first electronic Matriculation Examination for English as a foreign language is going to take place in the spring of 2018. Along with the digitalization of the exam format, foreign language examinations will be updated to include an assessment of oral skills which Finland has, on a nation-wide scale, lacked up until now, and the grades will be linked to the skill levels of the Central European Framework of Reference (CEFR, see Juurakko-Paavola and Takala, 2013). The first oral examinations are to be held in 2019.

The transformation of the exam walks hand in hand with the development of teaching technology in upper secondary language education. Changes to the examination have been called for because it no longer reflects the reality of language classrooms. For example, essays are no longer written by hand as in the exam but on the computer where they are easily editable and can be electronically shared with the teacher to reduce paper costs and ensure easy archiving. Secondly, the importance of oral skills has been stressed to the extent that an oral skills course was added to the foreign language curriculum in 2010, and yet the examination does not measure the student’s oral capacities. Students and teachers use more and more ICT in the everyday language education, but there is no way for students to display subject-specific ICT skills in the Matriculation Examination.

Certainly, many schools have long been utilizing ICT efficiently and regularly both in teaching and testing. Finnish upper secondary schools have, however, thus far had very diverse policies and levels of equipment when it comes to using ICT (Anttonen, 2011; Suomen Lukiolaisten Liitto, 2013). If changing classroom practices call for changes in the Matriculation Examination, the reverse then also holds true. The transformation of the examination inevitably leads to
significant changes in the upper secondary language teaching practice, especially in schools which hitherto have not invested in enhancing the use of ICT in teaching. The final courses of the upper secondary school language curriculum lean heavily on preparing the student for the Matriculation Examination and aim to give the student the best possible tools to succeed in the exams. In addition to learning the required language skills, the student needs to become comfortable with the format of the test as well as learn the requested answering technique. Traditionally, these have been practiced through using task types that imitate those of the exam all throughout the upper secondary language curriculum, featured in textbooks and other learning material used in class. This means that changing the exam will bring changes to classroom practice and learning material.

A perspicacious reader might already observe a certain vicious circle forming. The examination has a central position in defining the activities taking place in the upper secondary school language classrooms. Yet the aim of the examination is in essence to determine how well the student has achieved the goals of the curriculum. The examination therefore both assesses what has been learned and simultaneously defines what is to be learned, how and in what pace. It is therefore crucial that the examination with this double role is relevant and measures language capacity in a meaningful way.

The weakest voice in the debates around the reform has thus far unfortunately been that of the upper secondary school student participating in the examination. Although vivid discussion is taking place in social media, teacher forums and educational events, this discussion is mostly between teachers and professionals in education. The Digabi project works together with teachers and educational specialists throughout the nation, but organized collection of student opinion has thus far been scarce. This study aims to bridge the gap between English students, teachers and decision makers and find out how upper secondary English students view the change of format. Another goal is to find out what changes, if any, they feel are necessary in upper secondary English classroom practice as a result of this reform.

This study was initiated at a point when it first became possible to test-run the new task types through experimenting with the realization of example exercises from the digitalized Matriculation Examination, published by Digabi in the winter of 2014. These examples are in Finnish and applicable for any language. They represent the future of the digitalized Matriculation Examination for foreign languages, some more
ambitious and revolutionary in the way they utilize the possibilities of the electronic format than others. What was considered specifically important in the research design was that the students’ answers were not solely based on individual attitudes and previous experiences on language learning but on the experience of completing concrete example exercises from the exam.

The next chapter will provide theoretical background for the research topic, dealing first with issues related to language learning and then those connected to language testing. Chapter 3 will describe the research design, lay out the research questions and provide information about the participants. The results of the study will be discussed in chapter 4, which, along with illustrating the data and presenting an analysis of the participants’ answers, offers an abundance of quotations from the students themselves.
2 Theoretical background

As one of the focuses of this study is the relationship between the new electronic Matriculation Examination and current English teaching in Finnish upper secondary schools, it will touch upon the fields of both computer-assisted language learning (CALL) and computer-based language testing (CBLT). Since teaching and testing practices are inherently related, I will begin with a literature review of past and ongoing research on CALL (section 2.1) to provide a better understanding of the basis for the following discussion of CBLT (section 2.2).

2.1 Computer-assisted language learning (CALL)

This section is divided into four subsections. First, 2.1.1 will provide some basic background information regarding CALL research. Subsection 2.1.2 will then consider the benefits supporting CALL application, after which 2.1.3 will look at the methodology and devices related to CALL. Finally, subsection 2.1.4 will contextualize the discussion by examining Finnish research related to the phenomenon.

2.1.1 History and background for CALL

21st century language learning is immensely affected by technology. However we choose to refer to the devices – be it computer, smartphone, tablet or, simply and broadly said, technology – and however we choose to name its role in learning – whether it acts to assist, enhance or serve as basis for learning – its presence and effect are central in the present-day language classroom. In order not to be lost in the jungle of available terminology (e.g. TELL for technology-enhanced language learning, CASLA for computer-assisted second language acquisition, NBLT for network-based language learning, etc.) this study will employ the broadly used term CALL for computer-assisted language learning as a broad umbrella term including all possible variations where technological devices are used to assist or enhance language learning.

Thomas, Reinders and Warschauer (2013: 6) characterize the development of CALL as “a slow and uneven change process” and avoid using words like “revolution” or “transformation” to describe it. It has evolved along with the

---

1 For a more detailed analysis on terminology concerning CALL, see Levy (2005). For a useful glossary on the acronyms in the field of both CALL and CBLT, see Brown (2013: 75-76).
development of computer technology, with devices evolving from large, immovable computer installations to portable devices with wireless internet available almost everywhere (Davies, Otto and Rüschoff, 2013). It is evident that the implications on language learning of the former will hardly resemble those of the latter and what is relevant today may well be obsolete as early as a couple of years from now.

Although language is, in its very nature, a constantly changing and evolving phenomenon, theories of learning and teaching have often been introduced as lasting truths. Such truths are impossible in the field of contemporary CALL, where the assisting technology is constantly changing and its utilization hence requires constant updating of both pedagogical and technical knowledge concerning their usage. This of course adds to the workload of teachers, some of whom perceive technology as a threat both to their freedom and independence in the classroom and potentially even to their future employment. Echoing Clifford (1987), Blake (2013: 14) predicts that while computers will not replace teachers, “teachers who use technology will probably replace teachers who do not”. Using technology in language learning is no longer a choice made by the teacher but it is becoming a quality of education mandated by individual schools as well as the new Finnish National Curriculum².

Multiple attempts have been made to create specific categories within CALL, whether historically, pedagogically or theoretically. Warshauer and Healey (1998) proposed the historically aligned categories of “Behavioristic CALL”, “Communicative CALL” and “Integrative CALL”. Bax (2003: 21) argues against this categorization, accusing it of uncleanness and excessive linearity, and introduces his own categorization of approaches to CALL: “Restricted CALL”, “Open CALL” and “Integrated CALL”. In the first two categories, CALL is an “optional extra” (21) and it is not integrated into the syllabus. The final model is one where CALL is normalized, “embedded in everyday practice” (23), and used as a tool for learning in every lesson to promote interaction, interpretation and evaluation. Bax set this model as the goal in 2003. A decade later, Walker and White (2013) acknowledge Bax’s categorization, but make the addition of TELL (technology-enhanced language learning), suggesting a type of evolution from the early approaches of CALL to, finally, TELL. They argue that the state of normalization already pursued by Bax’s

---

² The new curricula for Basic and Upper Secondary Education in Finland are to be published in 2016; see National Board of Education (2014) for the draft for basic education.
stage “Integrative CALL” is only achieved in TELL. For a table depicting the approaches as put forward by Walker and White, see Appendix 1.

When revisiting the concept of normalization and integrated CALL in his 2011 essay, Bax introduces practical steps on how to carry out the process of normalizing technology in the classroom. These steps are 1) a “Needs Audit” which asks whether the technology in question is necessary or useful in the classroom, 2) devising a “Learning Plan” acting as a structured plan for the implementation of this new technology in a language education setting, 3) a context-sensitive “Research Programme” to be run concurrently with the previous step, aiming to spell out what, in that specific learning context, hinders or endorses the normalization of technology (12-13).

In 2011, Bax still sees the need for these steps eight years after his widely acknowledged article underlining the importance of CALL integration in 2003. This demonstrates how the process of proper and complete integration of a new learning tool into classroom practice, making it a normalized component thereof, demands plenty of time and research (Fulcher, 1999). In his discussion of the third step, Bax (2011: 12) expresses the same concern as brought out by Huh and Hu (2005) and O’Dowd (2007): that CALL research is excessively focused on theory-based study, largely ignoring the context- and culture-based variables which affect the implementation of these theories. In order to support those working with creating the “Research Programme” of Bax’s third step, he calls for more context-bound qualitative research concerning CALL. The present study attempts to make up for this lack by looking at the changing upper secondary school testing procedure and combining contemporary CALL and CBLT theory with information about the specific Finnish educational context. The current educational context in Finland, with big curricular changes to be taking place in the next few years and with the acceleration of making ICT an integral part of teaching and learning, provides particularly fruitful grounds for this type of study.

2.1.2 Why use CALL?

What, then, is so crucial about ICT that it deserves this integration into present-day language teaching? Aspects of ICT have been part of Finnish schools for decades: over-head projectors, nowadays often replaced by document cameras and/or interactive white boards, along with videos and other audiovisual material serve as
important tools for teachers to display learning material and computer laboratories are important for projects and other special schoolwork. As described above in 2.1.1, however, the furthest stages of CALL integration suggest learning technology where the student is an active and autonomous user of digital devices and ICT. Often arising assets accompanying this type of language learning are increasing authenticity, engaging the learner as a whole, encouraging active participation of the student, and, finally, the central position of ICT in today’s language use (e.g. Blake, 2013; Thomas, Reinders and Warschauer, 2013; Walker and White, 2013).

To begin with, augmenting the authenticity of learning material has been the feature perhaps most praised when it comes to learning languages through technology. Blake (2013: 4), for instance, characterizes the Internet as the “next best alternative to actually going abroad”. The Internet provides the learner with access to an authentic environment which native speakers of the foreign language in question also use daily. Moreover, it enables establishing relations with speakers of this language, both native and non-native, and enriches possibilities for communication and self-expression in this language (e.g. Chapelle, 2001).

Secondly, enhancing language teaching with the use of digital media has been promoted by highlighting its ability to engage the learner as a whole. Thomas, Reinders and Warschauer (2013: 2) underline the importance of learning “with digital media rather than merely via digital media” (emphasis in the original). This kind of learning is supported by Papert and Harel’s (1991) constructionist theory according to which best learning takes place when learners are involved in a making process. In fact, Thomas, Reinders and Warschauer add that Web 2.0-based technologies make learners “increasingly able to produce as well as merely consume content and learning materials”, using services like blogs, wikis and social network sites (2013: 4, emphasis original; see also Warschauer and Grimes, 2007). Such production of learning material acts to engage the learner in diverse making processes and has desirable effects on motivation (Grimes and Warschauer, 2008). Technology can have multiple roles in language learning3, which demonstrates its ability to engage the learner in diverse ways.

Encouraging the active participation of the student is another reported asset of ICT-enhanced teaching. Research has reported rising interest towards homework: as

---

3 See Stevenson (2008) and Walker and White (2013) for discussions on the different roles of technology in language learning.
much as 75% of the students interviewed by Grimes and Warschauer (2008) said that they spent more time on homework since the beginning of a one-to-one laptop program. One-to-one laptop programs, where each student has their own laptop at use, have been shown to particularly increase the students’ active participation and autonomy (see 2.1.3. for more discussion on 1:1 computing). Walker and White (2013: 13) believe that effective use of technology to support language learning helps learners become “more independent, collaborative, and engaged”. It can therefore act effectively both to support the learner’s independence and to encourage group work and collaboration between students. In addition, Suomen Lukiolaisten Liitto (2013) shows that using digital learning material increases overall student interest and motivation.

However, as Anttonen (2011: 7) states, it is difficult to claim that CALL is in essence a better or more efficient way to learn a language than traditional methods which do not make extensive use of technological devices. Using technology in language learning improves aspects of language learning, but the importance of its integration into classroom practice is based on changing language use. Communicative competence and computer literacy make for a new, integral part of functional language capacity, and encouraging the development of these skills has become an additional target for language teaching and learning (Chapelle, 2001). This makes it the duty of all teachers, arguably especially of language teachers, who carry a major responsibility when it comes to developing communicative skills.

2.1.3 CALL methodology and devices

When the implementation of CALL is concerned, two issues which often become stumbling blocks are methodology and devices. The importance of suitable teaching methodologies for the use of CALL has been recognized widely, e.g. by Blake (2013) and Walker and White (2013), while appropriate alternatives for device ratios and usage models are being discussed and researched (e.g. Bebell and O’Dwyer, 2010; Hatakka, Andersson and Grönlund, 2013).

To begin with methodology, Anttonen (2011) reported insufficient or inexistent institutional guidelines on the subject-specific implementation of CALL in Finnish upper and lower secondary schools. The participating teachers generally found technological and pedagogical support useful and even necessary, but only a half of the teachers found the received technological support adequate. For
pedagogical support, this percentage was less than half. According to Bax (2011), what impedes the normalization of technology in the language classroom is the poor design of attempts to bring technology into an educational institute. A concrete example of this in Finland is that many schools in the recent years have procured interactive white boards, but teachers have not been familiarized with their usage. This has merely led to the presence of a new device in the classroom, which, in itself, does little to promote any kind of learning or activity – on the contrary, inefficient use of newly acquired devices, such as interactive white boards, demotivates and frustrates learners (Suomen Lukiolaisten Liitto, 2013).

For understanding the theory behind utilizing technology for language learning as well as for practical suggestions for its realization, Walker and White (2013) provide a useful teachers’ guide. They suggest that rather than learning to use each newly available technological tool separately, teachers should have the chance to develop a fuller understanding of how and why to use ICT to support language learning. As a result of this kind of understanding, “any new use of technology which appears on the horizon can be integrated into teaching and learning in a principled way” (xiii). Naturally, both the acquisition of this kind of understanding and reviewing and updating the material and activities to be used in class requires resources directed towards continuing teacher training.

When it comes to devices, a rapidly growing trend is the aforementioned 1:1 computing. It is increasing in popularity in Finland, and it is a realistic view of future classroom technology. Bebell and O’Dwyer (2010) predict that this model in some form will be the future norm for American classrooms, and the reform of the Matriculation Examination has acted as a push from policy makers to accelerate this process in Finland, too. It is a logical implication that since the Matriculation Examination will be a 1:1 laptop test, the same model should be brought to the classroom. According to Suomen Lukiolaisten Liitto (2013), students would generally be willing to purchase a device, for example a laptop, to be used at school. Though this development is only in its infancy, many Finnish upper secondary schools have already encouraged new students at their school to buy their own laptops starting from the fall of 2014, and several primary or lower secondary schools have received funding for e.g. tablets for each student.

Case studies on 1:1 laptop usage report both positive and negative sides to the new kind of learning (see Bebell and O’Dwyer, 2010; Grimes and Warschauer, 2008;
Hatakka, Andersson and Grönlund, 2013). For students, positive effects include an improved and more fun learning environment, increased length of students’ essays, better scientific skills, deepened student engagement, possibilities for self-paced learning, improved critical thinking as well as more diverse use of sources. Teachers reported that managing and administrating the courses became easier and reading, assessing and giving feedback on essays became faster and could be reinforced with the help of evaluation software.

Negative aspects revealed in these case studies are the disturbing use of social media, increased plagiarism and difficulties with assessing online information. Participants of Hatakka, Andersson and Grönlund (2013: 108) felt that the laptop was a “distraction that diverted their attention from the learning and that they lacked strategies on how to overcome this distraction”. Other drawbacks of 1:1 usage mentioned by students were becoming less active, replacing face-to-face communication with online communication, the over-dependence of teaching and learning on computers and health issues such as reported back problems and headaches. In their synthesis of four empirical studies concerning 1:1 computing, Bebell and O’Dwyer (2013) found that the responsibility of the implementation often falls excessively on the teacher, which makes the teacher the biggest factor in the success or failure of such a project. This finding supports what was stated in the above discussion: the support received by the teacher is essential.

2.1.4 The use of ICT in Finnish upper secondary schools

The Finnish educational system has received praise in the last decades, but in the area of ICT-enhanced education, Finland is seen to be lacking behind other countries (Suomen Lukiolaisten Liitto, 2013). In the past few years, though, teaching technology has been a frequently discussed theme in the field of education. Studies focusing on language education specifically still remain scarce. Themes which arise from what CALL-related research exists are differences in ICT equipment levels and policies, lacking guidelines and deficiencies in both teachers’ and students’ ICT skills.

Both Anttonen (2011) and Suomen Lukiolaisten Liitto (2013), who interviewed teachers and students respectively, show substantial differences in the ICT equipment levels in Finnish upper secondary schools. According to Anttonen, tools available for teachers vary significantly both in amount and quality, and
teachers are often unsure of how to use or utilize them. Computers available for students range from none to a ratio of almost 1:1 (Suomen Lukiolaisten Liitto, 2013).

In schools of the latter ratio, students are encouraged as a policy to bring their own device to school, whereas other schools prohibit it. The majority among the students who answered the Suomen Lukiolaisten Liitto (2013) survey felt that ICT was not adequately made use of at their school and hoped for more. Some saw it as an inevitable part of future education and general knowledge, some considered it beneficial for their own specific learning style and motivation, and some saw internet-based material as offering altogether better chances for learning. They considered that the use of ICT equipment was inefficiently guided and they rarely trusted the teachers’ mastery of its use. These claims support the findings of Hurme, Nummenmaa and Lehtinen (2013), where a fifth of the participating students had never made use of a portable device at their own school, even if the school had ones to offer. It is easy therefore to agree with Anttonen’s (2011) conclusion that vague institutional guidelines concerning ICT usage have led to vague and ineffective practice.

The findings of Suomen Lukiolaisten Liitto (2013) imply that there is little if any discussion between teachers and students regarding the device purchases and the students do not feel empowered to affect these choices. Suomen Lukiolaisten Liitto suggest a change in this type of culture: making the whole school community part of the discussion around IT-related purchases would be advantageous because it would make each party responsible of taking care of the new equipment and observing the rules concerning its usage, agreed upon together. This would be an excellent basis for combining the know-how of teachers and students in the everyday classroom ICT usage and also for relevant use of resources.

Policies differ considerably not only between upper secondary schools, but also between teachers within a school. Some teachers have already adopted ICT and digital learning material into a natural part of schoolwork, whereas others’ use of it is incidental or nonexistent (Suomen Lukiolaisten Liitto, 2013). According to the students interviewed by Lakkala and Ilomäki (2013), out of all subjects, ICT is used the most in foreign language lessons. Nevertheless, the English teachers participating in Anttonen’s (2011) study were often uncertain about their own abilities using the available technology. The teachers in Lakkala and Ilomäki’s (2013) case study additionally report that since limited ICT resources restrict the use of e.g. computer
laboratories to few classes at a time, planning teaching where ICT plays a central role is difficult. In addition, teachers who were interviewed brought out the effect of big group sizes: few classes can provide a 1:1 laptop ratio to more than 30 students.

The rapid pace of the development towards computer-assisted learning in schools makes it difficult for teachers to keep up, but they are willing to develop their use of ICT for teaching. Although Anttonen (2011) reported hesitant attitudes of English teachers in Finland towards the benefits of CALL, many of the same teachers see themselves using ICT in their teaching in the future: 89% of the teachers agreed to some extent that ICT was part of their future plans. Kankaanrinta, Mikkonen and Vähähyppä (2012: 10) looked at the views of all teachers in Finnish lower and upper secondary schools, and 64% of them would be willing to use more ICT in teaching. 69% of them had perceived a change in the past few years in their way or amount of using ICT in teaching, and in the two years since this study was conducted, this percentage can only have grown. Teachers are becoming more and more open to enhancing their teaching with ICT. At the very least, they feel pressure to do so, regardless of their hesitance.

In addition to the teachers’ imperfect digital skills, several studies concerning Finnish students’ ICT abilities show that even students lack in relevant IT skills as well as critical research skills vital to higher education (see Hautamäki et al., 2012; Kiili, 2012; Pajarre, 2012). According Lakkala and Ilomäki (2013), teachers and educational decision-makers have perceived digital study skills as little more than basic IT skills which students master through informal learning taking place in their free time, but the above studies suggest the opposite. Lakkala and Ilomäki come to the conclusion that upper secondary education, in its current form, fails to develop relevant modern-day skills. They believe that the Matriculation Examination reform will produce pressure to both develop vital ICT skills and utilize new kinds of working methods essential to both the new Matriculation Examination and further life and studies.

2.2 Computer-based language testing (CBLT)

This section will go through some history and background for computer-based language testing (subsection 2.2.1), then look at a few example WBLT realizations in 2.2.2 and finally assess the effects of the change in format on the Matriculation
Examination, basing this discussion on Bachman and Palmer’s (1996) test usefulness qualities (subsection 2.2.3).

2.2.1 History and background for CBLT

While CALL has remained to this date a frequently used acronym for computer-assisted language learning, even though other acronyms have been suggested as better-suited, there is more variety in terminology within the field of language testing. CALT is used by many (e.g. Chapelle, 2001) to refer to Computer-Assisted Language Testing, but this ambiguous acronym is also used to denote Computer-Adaptive Language Testing (e.g. by Brown, 2013), which is a specific form of CBLT. CBLT will therefore be used in this study because of its clarity and suitability when it comes to the Finnish electronic Matriculation Examination, which, in contrast with language learning and teaching, will be computer-based rather than computer-assisted.

Brown (1997 and 2013) has produced two excellent overviews of the history of CBLT literature. In 1997, Brown found that CBLT literature was concerned with item banking (“any procedures that are used to create, pilot, analyze, store, manage, and select test items so that multiple test forms can be created from subsets of the total “bank” of items” (44)), how to use the newly available technologies, prospects and challenges related to CALT and the effectiveness of computers in general for the purposes of language testing.

In addition to dealing with some basic existential questions, the earlier topics show that since the beginnings of CBLT research, the prospect of computer-adaptive language tests, CALTs, has aroused a lot of interest in researchers. In the electronic mode, it becomes possible to create an intelligent test which determines the test-taker’s level and afterwards begins to present him only with tasks which correspond to this level. The software can, for example, begin with tasks at the middle level and then begin to adapt to the test-taker’s level: if he succeeds in the task, it will produce an exercise of a higher level, whereas if he cannot complete the exercise, he will next face an easier task. Alternatively, the test can begin with easy exercises and become more and more difficult, stopping at the student’s own skill level. The software therefore constantly re-assesses the student’s level and presents him with tasks surrounding this level. Each test-taker can then complete a tailor-made exam defined according to his skills.
This type of testing has clear assets when it comes to the test-taker’s language learning in the test situation: no one needs to struggle to finish an exercise which is clearly above his level, nor must he spend time on tasks which are too easy and do not develop or challenge his language skills in any meaningful way. CALT offers each test-taker an individual test experience specific to his own level. For more theory on CALT, see for example Chapelle (2001).

As for the usability and relevance of this kind of testing measure in Finland, a report was published by the Matriculation Examination Board examining the possibilities of CALT concerning the Matriculation Examination (Hakola & Mäenpää, 2013). It is argued in their study that the adaptive test measure gives a fuller account of the test-taker’s ability rather than determining the level at which he can accomplish a particular, prearranged and limited set of tasks designed to measure this ability. This type of measuring, of course, requires a massive bank of test items for each difficulty level and would therefore be challenging in the context of the Matriculation Examination taking place twice a year and traditionally excluding the chance of using the same task more than once. Additionally, no research was found on the adaptive measurement of oral proficiency. Reliable assessment of the student’s own written production in an adaptive test is another challenge to which few solutions were offered. Consequently, the adaptive test measure will be unlikely to be used in the Matriculation Examination for English in the present conditions. It is, however, a prospect rendered possible by the electronic format and could be used to assess some part of the student’s language capability in the Matriculation Examination for English in the future.

Excluding, then, the topics dealing first and foremost with the adaptive test measure, Brown (2013) found that resurging topics include CBLT content (assessment of different areas of language skills using CBLT), example CBLT realizations and the tools and resources available for those interested in CBLT. Finally, Brown introduces a broad label “CBLT delivery issues” under which he bulks the discussion of diverse issues such as those concerned with CBLT/PBLT (for paper-based language testing) comparisons, WBLT (for web-based language testing), interface architecture, test-takers’ experiences and teacher training. Out of these diverse topics being discussed in the field, we will look at those most relevant to our topic of the reform of the Finnish Matriculation Examination, namely CBLT/PBLT
comparisons, example realizations of WBLT and test-taker experiences brought out in these realizations.\(^4\)

### 2.2.2 Why use CBLT?

In order for an educational institution to demolish a functioning PBLT system and introduce a replacement in the form of CBLT, strong gains need to be attainable by the reform. Some reported advantages of CBLT compared to PBLT are the ability to store answers securely on the computer, teachers’ easy access to answers regardless of time and space, possibility for automatic and instantaneous assessment as well as feedback, and the increased speed of editing and creating longer answers such as essays using a keyboard (Alderson, 2000; Laakso, Kaila & Rajala 2014). The electronic format may relinquish both the test taker and the assessor from the restrictions of time and location. Another advantage especially relevant in the Finnish upper secondary school context is the positive washback on teaching and learning, leading to an increase in the interest towards ICT-assisted classroom activity (Lakkola and Ilomäki, 2013).

As disadvantages to CBLT, Alderson (2000) mentions the limitedness of machine-scorable task types, disabling for example the assessment of writing and speaking in meaningful ways, and the degree of computer literacy required of the test-takers. He states that “the need to move to and fro through screens is much more limiting than being able to flick back and forth through print” (594). This is decreasingly true to today’s language learners and even more to those of tomorrow. In their free time, they are becoming more and more comfortable with reading text in the electronic format and operating with multiple windows and screens simultaneously. Searching for information in various sources is, for many, more natural and convenient in the electronic mode, and to today’s students, it is already a more familiar way of working than consulting several paper sources. In this respect, no “limitedness” of form is found in current computer-based language tests with regards to the forms utilized in teaching and present in language use.

CBLTs and PBLTs are, in fact, governed by different limitations. On the one hand, a PBLT is much more limiting and limited than a CBLT, for instance when considering the editability of the exam, amount of source material that can be made

---

\(^4\) To the interested reader, Brown (2013) offers an abundance of references to noteworthy works on all of the above topics.
available to the student, and the modality of this material (e.g. audiovisual material). On the other hand, a CBLT is limited by constraints regarding the test software, specific program functions and the level of the student’s basic ICT knowledge. The final demand will be further dealt with in the discussion on validity in section 2.2.3.

The disadvantages reported on CBLT often stem from the expectation that PBLTs and CBLTs should produce identical scores and the format should not affect the test results in any way. This can be considered natural, as CBLTs are often digitalized versions of what used to be PBLTs. Fulcher (1999: 292) asserts that any attempt to achieve equivalence of forms is likely to fail. Scores will vary between the pencil-and-paper and CBT forms of a test. However, it is still incumbent upon an institution to ensure that significant score variations associated with a CBT do not introduce a bias into the results.

According to Fulcher, then, it is natural that a change in the format of the test affects score distributions. This, in itself, is not a problem for him. It only becomes a problem if a bias is introduced into the results, for example favoring certain types of socio-economic backgrounds or weakening the position of students of particular linguistic backgrounds.

2.2.3 Example WBLT realizations

The Finnish Matriculation Examination will be a web-based language test, a WBLT, with online access only to the Matriculation Examination Board’s own server. Free internet access is therefore denied to test-takers, though the possibility of allowing it in the future is being discussed.

To take an example of a WBLT-related study, Fulcher (1999) conducted a case study experimenting with computerizing a university-level English language placement test consisting of 80 multiple-choice items and two essays. The interest of the study was primarily to find a way to reduce the workload of teachers: in the context of the study, teachers had to grade up to 120 students’ papers within one day. Although the essays still had to be graded by the staff even when using the WBLT format, the automation of multiple-choice answers was seen to lighten the workload considerably. Since the number of test-takers was great, Fulcher chose WBLT to minimize the need for technological support in the installation of special testing software.

Fulcher found that CBLT was “sufficiently reliable for its purpose” and correlated with PBLT, though “not highly enough for accurate prediction of a score
on one form of the test from a score on the other form” (1999: 298). Furthermore, it was discovered that the CBT was “more sensitive to variation in language ability”, which was seen to lead to a more accurate placement decision.

The context is similar to that of the paper-based Finnish Matriculation Examination: currently, teachers work on a tight schedule to mark and grade a large number of answers, including ones which could be entirely machine-gradable. In the WBLT solution, a considerable number of test-takers are completing the exam on individual computers. The one significant difference between Fulcher’s placement test and the Matriculation Examination is that the placement test is a relatively low-stakes test. If a student is placed into a group below his level, he can be moved onto a higher-level group. This type of correction cannot be conducted when it comes to Matriculation Examination grades (improving a grade is only possible through retaking the exam, which can only be done once) and therefore students who achieve worse results simply based on the format are at a more serious disadvantage. In contrast, the increased sensitivity of the exam to varying levels of language ability enables more accurate grading and can be regarded as a significant gain.

As another example realization of WBLT, a Finnish study by Laakso, Kaila and Rajala (2014) briefly describes a case study in which the virtual learning environment ViLLE was used by upper secondary students in two schools for a preliminary examination for English aiming to prepare the students for the Matriculation Examination and imitating its conditions. The students had the opportunity to choose whether to complete the preliminary exam on paper or on the computer. After completing the preliminary exam, the students answered a survey with statements to be graded on a Likert scale. These students were highly content with the exam arrangements: they unanimously considered the experiment successful and, given the choice, the students would, with one accord, rather complete the actual Matriculation Examination in the digital mode utilizing the environment in question than on paper. Another finding was that students spent relatively little time on the exam: the average time used for it was 2 hours 27 minutes, even though 6 hours of exam time was provided (equivalent to the exam time in the Matriculation Examination).

The goal of their survey was, however, more to concentrate on the usability of the software rather than the update of the examination, as the exercises used in the preliminary exam were Matriculation Examination exercises from previous years.
which had been converted into the electronic mode without further modification. The findings are still relevant to this study as the students in question felt that this electronic surrounding provided them with an equally suitable platform for bringing out their language skills as the paper-and-pencil format, if not even a more suitable one.

2.2.4 CBLT for the Finnish Matriculation Examination

When the digitalization of the Matriculation Examination became a potential option, the National Board of Education along with the Matriculation Examination Board began following similar projects around Nordic countries, the EU region and worldwide. It was found that relatively few EU countries had systematic high-stakes examinations in the electronic format (Lahti et al., 2013). Influences for the digitalization of the Matriculation Exam have mostly been drawn from Denmark, where an electronic upper secondary school final exam is fully functioning, along with Norway, Holland, Hungary and Romania (Lahti et al., 2013). Ideas for possible future implementation of CALT have been drawn especially from Iceland, Denmark, Holland and Great Britain (Hakala and Mäenpää, 2013).

Some studies already exist on the attitudes of upper secondary school students and teachers concerning the use of ICT in testing, and even a few dealing especially with the Matriculation Examination. However, these studies do not specifically concern English or foreign languages. One such is Hurme, Nummenmaa and Lehtinen’s (2013) report published by the National Board of Education, where some 40% of the students wanted to make use of a computer in the Matriculation Examination for foreign languages. In contrast, 37% would only resort to traditional tools. First-year students were found more willing overall to utilize technology in the exam. Perhaps surprisingly, experience on electronic course exams was not seen to affect attitudes concerning the digital Matriculation Examination, but those with strong IT skills preferred the digitalized version of the Matriculation Examination in all other subjects than mathematics. In the survey conducted by Suomen Lukiolaisten Liitto (2013), some students expressed their concern about the Matriculation Examination turning electronic, suggesting that the use of ICT should be improved in the schools before this happens.

Interestingly, Lakkala and Ilomäki (2013) found substantial differences in the ways teachers and students estimated the students’ performance in imagined
Matriculation Examination exercises. These exercises included e.g. scanning multiple online sources of information and forming a synthesis, answering complex questions in a foreign-language test with the help of an online dictionary and evaluating the correctness of given information based on information that can be found online. Hardly any teachers stated that students would do well in these exercises, whereas the students were more confident about their own skills. Lakkala and Ilomäki suggest that this demonstrates a challenge for both teachers and students: Teachers need to learn the pedagogy behind teaching these kinds of skills with the help of ICT, whereas students need to work on recognizing their own abilities in relation to the skills that completing the exercise requires.

The shortage of the above studies when it comes to the Matriculation Examination, however, is the fact that the students’ answers were solely based on their own attitudes and ideas concerning the new Matriculation Examination, as nothing concrete could yet be offered to them in order for them to be able to answer based on experience. Students participating in the current study were able to base their answers on the experience of completing actual example exercises and did not have to rely on hypotheses.

2.2.5 Test usefulness qualities

In this final sub-section dealing with CBLT theory, we will review the basic principles for any language test development as put forward by Bachman and Palmer (1996) and see how each of these features is preserved, enhanced or diminished in the new exam format (much in the style of Chapelle 2001:Ch.4). These features will reappear in the results section of this study and were considered crucial to revisit in order to be able to evaluate the effects of the reform on these fundamental test features.

Bachman and Palmer (1996) base their approach to language test development on two “fundamental principles”:

– The need for a correspondence between language test performance and language use:

In order for a particular language test to be useful for its intended purposes, test performance must correspond in demonstrable ways to language use in non-test situations.

Chapelle’s (2001) examples are specific example tasks within a particular testing context, which enables her to closely analyze the related test usefulness criteria. As the Matriculation Examination consists of several task types and measures overall linguistic capability, this type of specific analysis is not possible here. Only the form and style of Chapelle are replicated or imitated in this discussion.
A clear and explicit definition of the qualities of test usefulness:

Test usefulness, consisting of several qualities (reliability, construct validity, authenticity, interactiveness, impact, and practicality), is an overriding consideration for quality control throughout the process of designing, developing, and using a particular language test. (9, emphasis original)

The first principle essentially requires the exam results to reflect existing language skills. If a student does well in an exam, that should mean he is capable of handling a corresponding real-life language-use situation. This is a broad and fundamental principle and requires many of the qualities to be discussed below in order to be fulfilled.

The second principle consists of six separate qualities. Chapelle (2001) argues that a test cannot fulfill all of these qualities, because any attempt to maximize one is sure to diminish one of the others. For example, creating an exam task as authentic as possible would probably make it quite impractical and therefore difficult to realize. Hence, any test can only illustrate these qualities to a certain extent. In light of the purpose of the test, its planners should consider all the qualities and be as transparent as possible about recognized deficiencies thereof.

2.2.5.1 Reliability

The first of these test usefulness qualities is reliability. Bachman and Palmer define this as “consistency of measurement”, and it is also listed in Weir’s (1993) principles for good test design. According to Bachman and Palmer’s explanation, test A should differ from test A’ “only in incidental ways” (20). The Matriculation Examination has long traditions and it is natural that English exams have evolved along with the introduction of new learning and testing theories and the increase in Finnish students’ English skills. Any test in the 2010s is then sure to be more difficult than tests in the 70s. The change in format is an important point in the development of the Matriculation Examination, and the goal is to make the shift as subtle as possible. Although the change of format is not an “incidental” difference, such a change is perceived as essential in preserving the relevance of the examination.

Bachman and Palmer list raters as a possible variable which can affect the reliability of the exam. Since the rating procedure will change considerably as it is brought to the electronic mode, the digitalization will bear consequences to the reliability of the Matriculation Examination. Multiple-choice items will be marked
instantaneously by the computer, which diminishes the chance for errors. In the
future, it will be possible to include more rating-related technology, such as speech
recognition or automatic essay scoring programs, to further enhance the reliability of
the exam. The digitalization also creates new challenges to the rating procedure,
since assessing the oral tasks requires the creation of new rating procedure specific to
evaluating oral performance.

2.2.5.2 Construct validity

Another quality of test usefulness mentioned by Bachman and Palmer is
construct validity, “the meaningfulness and appropriateness of the interpretations that
we make on the basis of test scores” (21). Chapelle (2001: 98) argues that validity for
CALT, or CBLT, in general cannot be defined, because validity “applies to test
interpretation and use in a particular situation”. This is why construct validity is
often examined through case studies. One such is that of Choi, Kim and Boo (2003),
who compared a paper-based version of a university-level English language
proficiency test with a computer-based version of the same test. Their findings
support to some extent that the same constructs are measured by CBLT and PBLT
and demonstrate an overall comparability of the two test scores, with highest
comparability rates for the grammar test and weakest for reading comprehension
(316).

The interpretation of the Matriculation Examination grade of English is
linked to the student’s assumed language ability as a whole: reading, writing,
listening and speaking skills. This grade, instead, affects the student’s possibilities
for further study and acts as an indicator of his language skills throughout these
studies and the beginning of working life. In the renovated exam format, all of the
four language skills will be tested in one exam situation for the first time. The exam
tasks are constructed to include more tasks where several skills are at use
simultaneously, and the oral dimension is added to the tasks. The new digitalized
Matriculation Examination should therefore be a better and more versatile indicator
of the test-taker’s language ability.

---

6 See e.g. Douglas and Hegelheimer (2007) for implications, experiences and analysis on speech
recognition and automatic essay scoring.

7 Despite this argument, issues regarding the construct validity of CBLT in general can be found for
example in Choi, Kim and Boo (2003: 297).
A fear concerning the construct validity of digital tests is that instead of measuring purely linguistic skills, a CBLT also measures ICT skills (Choi, Kim and Boo, 2003). It needs to be considered that when it comes to the Matriculation Examination, students are fairly homogeneous in comparison to the student material in large-scale high-stakes examinations like language proficiency tests (TOEFL, etc.). The basis for the exam is the upper secondary school curriculum, which currently acknowledges the importance of ICT (National Curriculum for Upper Secondary Education, 2003) and is likely to stress its importance even more in the future. After roughly 10 years of computer-based TOEFLs, Wall and Horák (2011, data collection conducted from 2006 to 2008) found that computational skills in the internet-based TOEFL had ceased to affect the test-taker’s performance. In Finland, the corresponding adjustment period is sure to be shorter. In order not to make the component of IT skills too big in a test which ultimately tests language skills, the software will be made as user-friendly as possible and, in addition, students will be familiarized with it before the exam situation\(^8\). The user-friendliness of the interface has also been recognized as a factor affecting construct validity by Fulcher (2003).

With the introduction of a new and sometimes unpredictable tool to be used in the test situation, the first years of implementation may well see a rise in test anxiety and computer anxiety, which can both affect the test performance. The central position of the Matriculation Examination with regards to the student’s future prospects already results, for some students, in test anxiety. Adding the oral component which demands the student’s immediate input may increase instances of test anxiety, as demonstrating oral capabilities has been seen to cause especial anxiety among language students and test-takers (for more on foreign language anxiety and test anxiety, see for example Horwitz, Horwitz and Cope, 1986; Gregersen and Horwitz, 2001; and Hewitt and Stephenson, 2012). Such anxiety is best prevented by making the situation as familiar to the student as possible beforehand, but cannot be eliminated altogether (e.g. Fulcher, 1999: 291).

2.2.5.3 Authenticity

Next, Bachman and Palmer’s test usefulness qualities bring us to authenticity. This criterion calls for a correspondence between the qualities of the test task and

\(^8\) The software will be published at the end of year 2014, giving the students plenty of time to become comfortable with it before the first digital exams in 2016. (Digabi, 2014)
those of a Target Language Use (TLU) task. Bachman and Palmer view authenticity as a vital test characteristic because it is closely related to construct validity: it “provides a means for investigating the extent to which score interpretations generalize beyond performance on the test to language use in the TLU domain, or to other similar nontest language use domains” (24).

As argued above in the section on CALL (section 2.1.2) this criterion is unquestionably better realized in the new digitalized format of the Matriculation Examination, and it is perhaps this principle that most importantly urges the change to take place. It has been recognized that students rarely find themselves in a non-test situation like the current Matriculation Examination. Almost twenty years ago, Russell and Haney (1997) expressed their concern about the increasing unfairness of testing essay writing in PBLTs, when learners are accustomed to composing longer texts on the computer. Since their work, a whole generation has grown up learning the culture of writing on the computer.

The situations where a foreign language is used in real-life situations occur increasingly in online environments: e-mailing a colleague in a foreign country, chatting online with friends from abroad, reading or commenting on blogs or news articles published online, holding video call conferences to other countries, etc. What is more, the updated Matriculation Examination will also be able to assess oral communication skills, which have become increasingly important for both working in globalized contexts and communicating with people with different linguistic backgrounds for other purposes.

2.2.5.4 Interactiveness

Interactiveness, the fourth quality of test usefulness, refers to the way in which the test engages the learner’s language ability as a whole. An interactive test task requires the student to combine his/her “areas of language knowledge, metacognitive strategies, topical knowledge and affective schemata” (25). In this respect, the new Matriculation Examination has the possibility of improving considerably, because it engages the test-taker as a digital learner and requires the use of more versatile language skills. Since the electronic format is capable of providing a lot more source material for the use of the student and is not bound by print expenses, it can engage the learner in new ways by including video, different text types, pictures, etc. The prospect of allowing free internet access to the test-taker
in the future would increase both the authenticity and interactiveness of the examination even more. These interactive communication tools have become increasingly available to the students in everyday language situations, regardless of their whereabouts, and they even enable interaction between test-takers during the exam as part of the test.

As the strategies at use in the electronic exam situation differ considerably from those used in the paper-and-pencil format, something of the interactiveness is also lost. It may be that learners who visualize for example their written answers with the help of mind maps or other hand-drawn figures before submitting them can feel at a loss when trying to transfer this process to a computerized test surrounding. This type of sketching by hand will be allowed in the digital examination, but ideally a plausible way to do this on the computer would be offered to the student. A visually oriented learner would also benefit considerably from the increased level of visual input in the source material rendered possible by the electronic mode.

On the other hand, students have also reported assets to writing by hand, such as deeper reflection on the topic at hand (Hatakka, Andersson and Grönlund, 2013). Handwriting may help some learners produce meaningful and carefully thought of essays. What is lost or gained in the writing process when the connection between the hand and the brain is transformed from producing a shape to tapping a key is highly subjective. This is an area calling for further research in the joint fields of learning psychology and language learning. Overall, in light of the present research it can still be stated that as far as interactiveness goes, there are more gains to the reform than losses.

2.2.5.5 Impact

A further quality involved in test usefulness is its impact “on society and educational systems and upon the individuals within those systems” (Bachman and Palmer, 1996: 29). This impact functions on two levels: a micro one which concerns the individuals, and a macro level dealing with institutions and society. As stated before, the impact of the Matriculation Examination on the individual is considerable as it affects his further opportunities. The positive impact on the individual is an often seen motivation rise at the end of upper secondary school, when further studies become a pressing question in the student’s mind. On the negative side, the examination causes both students and teachers stress and extra work hours.
In addition to the impact of the Matriculation Examination on the individual student, its impact on upper secondary school education is, as discussed in section 2.1.2, also substantial. The new digitalized format will and has already begun to have a massive impact on teaching in upper secondary education as well as the acquisitions and economy of schools and municipalities (e.g. Hurme, Nummenmaa and Lehtinen, 2013; Lakkola and Ilomäki, 2013). The washback of the exam on teaching and learning is great in both upper secondary schools and even earlier stages of education. This washback brings us back to CALL whose good and bad sides are dealt with in section 2.1. In the next few years, the change at hand is likely to lead to a period of rapid change and computerization of teaching and learning (Lakkola and Ilomäki, 2013; Suomen Lukiolaisten Liitto, 2013). After the reform, the Matriculation Examination diploma will have the added quality of serving as proof of basic computer literacy.

2.2.5.6 Practicality

The final test usefulness quality is practicality, which Bachman and Palmer define as “the relationship between the resources that will be required in the design, development, and use of the test and the resources that will be available for these activities” (36). These resources include 1) human resources, 2) material resources and 3) time.

Reducing costs is pursued when it comes to human resources and time. The digitalization will hopefully bring the “management and administrative efficiency” brought to the class by CALL to the Matriculation Examination Board, which, twice a year, conducts endless amounts of paperwork related to the exam (Hatakka, Andersson and Grönlund 2013: 95). Instead of having a separate exam session for the listening comprehension and the written part of the exam, the new language exam will be completed within the one session, which will reduce costs and inconveniences produced by several exam dates and locations. With the partially automated initial grading process, teachers save a considerable amount of time and the transportation of the results to the Matriculation Board will be almost instantaneous. Also the final stage of reporting, publishing and archiving the results will be more practical and less time-consuming using ICT.

Material resources will change entirely with the elimination of print expenses and the introduction of massive costs related to IT equipment and expertise. The
costs are not restricted to the one-time acquisition of the necessary equipment: the reform requires commitment to constant updating of the technology needed to complete the examination.

Overall, it is difficult to evaluate the practicality of the digitalized Matriculation Examination at this point of its development. Some aspects of practicality will suffer drawbacks whereas others are enhanced. The following years will surely require substantial amounts of resources directed towards technology which makes completing such a large-scale digital examination possible.

Appendix 2 shows a summary of the above discussion, similar to those of Chapelle (2001), showing the effect of change in format with regards to Bachman and Palmer’s (1996) test usefulness qualities.
3 Research design

The current chapter describes the design of this study. I will begin by presenting the research questions, then provide information about the participants and data collection procedures and describe the background for the example exercises used in the study. Finally, I will describe the design of the questionnaire along with a brief description of the pilot study.

3.1 Research questions

This study aims to find the answer to the following two research questions:

1. **Based on their experience and the example exercises, how do students perceive the digital Matriculation Examination for English?**
   - What kind of attitudes do the students convey towards the digitalization of the Matriculation Examination of English?
   - How do they describe their experience of the example exercises?
   - What are the students’ suggestions for improving the digital Matriculation Examination?

2. **In the students’ opinion, what is the relationship between the skills the example exercises are testing and the language education they are receiving in the upper secondary school?**
   - Do they feel the language education they are currently receiving prepares them for an exam like this? Should it?
   - What, if anything, do they think should be done differently in language education once the digitalized Matriculation Examination steps in?

The research questions will be answered with the help of questionnaire data gathered from upper secondary school students. Stronger emphasis will be laid on the first research question, since the data provided especially interesting viewpoints to its analysis.

3.2 Participants

The data consists of 93 second-year students’ answers to the questionnaire created for the purpose of answering the above research questions. The participants were aged 17 to 18 and studied in upper secondary schools around the Helsinki region. 39 (42%) of the participants were women and 54 (58%) were men. Second-
year students were selected as participants because generally at this point of their studies, the students have already been familiarized with the Matriculation Examination. They were also best equipped to answer questions concerning upper secondary English education, because they were the oldest students in the school at the time of the data collection (spring 2014) as most of the third-year students were already finished with their courses.

The participating students come from seven different schools around the Helsinki region (including the pilot group). Several English teachers of upper secondary schools around the metropolitan area were contacted via e-mail. In order to participate, the schools needed to provide a computer, headphones and a microphone for the use of each student and have a recording program installed on the computers. Many replies of teachers showed their uncertainty as to the availability of these appliances, some even declined because their IT class could not fit their entire group or because this equipment was unattainable. Several teachers also expressed their willingness to complete the experiment in all of their teaching groups, but the limited resources of the study only enabled one group from each upper secondary school to participate.

As participation was dependent on the activeness of the teacher, the results may be slightly inclined towards presenting the use of technology in classes higher than the average rate. Teachers who already utilize teaching technology may be more inclined to participate in this type of study, whereas teachers who generally do not might easily opt out of participating.

### 3.3 Data collection procedure

The data collection procedure was planned so that it could be completed within a 75-minute language class. The majority of upper secondary schools have 75-minute classes and the experiment was to be made as much like part of normal schoolwork as possible to ensure a relaxed and open atmosphere where the students could express their opinions freely. The first 10 minutes of the class were reserved for introduction and instructions, the next 45 for the completion of the example exercises, and the final 15 minutes for the questionnaire. An optimistic five minutes was spared for possible technical or other problems. Table 1 below demonstrates the data gathering procedure and its goals and refers to relevant material in the Appendices.
During the data gathering, both predictable and unexpected circumstances affected the test situation. It was to be expected that some students would arrive late, and in some groups, this amounted to as much as 10 minutes of lost time for the latecomers’ completion of the example exercises. In addition, one group with an unusually high amount of late arrivals also had to finish 15 minutes early due to an event which was not brought to the attention of the researcher or the teacher of the group beforehand. It is obvious that the difference in the time available for the questionnaire as well as the exercises affects their answers. Another factor which may have affected the students’ attitudes and hence the results of this study is the teacher’s attitude in the test situation. Some teachers would express their criticism towards the exercises and the test arrangements openly, whereas others remained impartial.

### Table 1 Data collection procedure

<table>
<thead>
<tr>
<th>Time</th>
<th>Action</th>
<th>Goal</th>
<th>Appendices</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 min</td>
<td>Introduction and instructions</td>
<td>- introducing the topic</td>
<td>See Appendix 3 for the instructions given to students (in Finnish)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- familiarizing the students with the tools used for the experiment</td>
<td></td>
</tr>
<tr>
<td>35 min</td>
<td>Example exercises 1, 2 and 3</td>
<td>- familiarizing the student with the current state of the new Matriculation Exam</td>
<td>See Appendix 4 for exercise scripts and screenshots (in both Finnish and English)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- giving them the chance to form an informed opinion</td>
<td></td>
</tr>
<tr>
<td>10 min</td>
<td>Example exercise 4</td>
<td>- giving the students an idea of what the oral exam might look like</td>
<td>See Appendix 5 for a script of the audio clip</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- collecting information for the development of the oral exam</td>
<td></td>
</tr>
<tr>
<td>15 min</td>
<td>Completing the questionnaire</td>
<td>- to answer the research questions proposed in section 3.1</td>
<td>See Appendix 6 for the questionnaire (in Finnish)</td>
</tr>
<tr>
<td>5 min</td>
<td>Reserved for technical problems, late arrivals and other unpredictable issues</td>
<td>- avoiding a sense of hurry and anxiousness both on the participants’ and the researcher’s part</td>
<td></td>
</tr>
</tbody>
</table>
3.4 Background for the example exercises

The example exercises used in the study were designed by the researcher, following the examples published by the Digabi project. They were considered crucial for the study, because without the experience of the example exercises, the participating students would have had next to no idea what the current state of development concerning the new language examination was and would have had to rely solely on the images they attach to digital exams and tasks. This study aimed to truly include the students in the process of developing the language Matriculation Exam, which was seen to require updating their information on the current state of its development.

The example exercises consist of: 1) one multiple-choice reading comprehension task, 2) another reading comprehension task with open questions in Finnish, 3) a vocabulary and grammar gap-filling exercise, and 4) an oral exercise testing both listening and speaking. The task types were chosen so that two of them (numbers 2 and 3) represented the traditional task types of the Matriculation Exam for English which were simply converted into the digital format with few adjustments. Even in these task types, though, there were improvements and additions made possible by the digital format. Such were for example being able to insert the answers into the text in the gap-filling exercise, allowing the students to better tell whether their answer fits or not, along with the possibility of adding a picture to support the content of the text.

Exercises number 1 and 4, then, represent more experimental and ambitious tasks in which the digital format plays an important role. The reading comprehension task 1 included a link to an outward source created on a Blogger platform, resembling a news site and including statistical elements as well as short articles. The questions required the students to familiarize themselves with all of the material and understand links between the articles. The oral exercise, on the other hand, was a task type rendered possible by the digital format: the students listened to instructions on an audio file and recorded their reactions simultaneously by using a sound recording program. In the first part of the task, the student was asked to read aloud a sentence visible on the screen. In the second part, he had to repeat a sentence he heard, next, choose between given alternatives, then retell a story in his own words and in the end, answer an open question asking his opinion. The audio instructions were recorded by the researcher, following closely the example published by the Digabi
project. The aim was to make the student transition smoothly from producing simple, restricted output to finally performing free speech. Appendix 4 and 5 contain screen captures and scripts for the example exercises.

As the focus of this study is on the students’ opinions, the answers to the exercises were not analyzed. Their purpose was to give the students a chance to form an informed opinion on the current state of the Matriculation Exam for English.

3.5 Questionnaire

The questionnaire contained both Likert statements and open-ended questions. It was designed with the help of Dörnyei (2003) and comprised of four main sections. The first one, labeled “The electronic Matriculation Examination for English”, included questions on the technical qualities and usability of the exercises as well as attitudinal questions concerning the new format. In the second section called “Comparing the paper version of the Matriculation Examination to the electronic version”, the students were asked to estimate their performance and experience in the example exercises compared to the paper versions they have previously completed for practice. The third section of the questionnaire was dubbed “The relationship between language education in upper secondary school and the electronic Matriculation Examination for English” and included questions regarding the current state and future of upper secondary language education in relation to this change. The final section, “Open questions”, included four open questions which were connected to the research questions. In addition to these four sections, there was a brief background information section asking about the student’s gender and asking the student to evaluate his English skills, English agreeability and IT skills on a school assessment scale used in Finland (10= excellent, 9= very good, 8= good, 7= satisfactory, 6= moderate, 5= adequate, 4= poor/fail). Because the instances of grades 4 to 6 were noticeably few compared to the other grades, these grades were combined into one group in the analysis to ensure adequate frequency levels for each grade set.

The data produced both qualitative and quantitative information and was analyzed using mixed-methods analysis to enable deep interpretative analysis alongside presenting moderately generalizable quantitative results (Dörnyei, 2003).
3.6 Pilot study

The study was piloted in an upper secondary school in the metropolitan area, with 22 pilot answers. Due to technical problems, the oral exercise of the pilot study was more than a half shorter than the final version used with the other groups. The IT class in question was designed for computer-based group work and therefore posed challenges when it came to conducting individual, silent work.

After the pilot study, the test instructions were revised to add emphasis on working silently so as to create a more test-like situation. As for the questionnaire, the options “I agree to some extent” and “I disagree to some extent” were deleted and a neutral option was added in order not to force an opinion out of the students on a topic they have little experience on. Some reported ambiguities in the example exercises were deleted.

The results of the pilot study do not differ considerably from those of the actual group. Similar suggestions came up and only a couple students commented on the short length of the oral exercise.
4 Results and discussion

The results chapter is divided into three main sections. The first section deals with the students’ experience of the digitalized Matriculation Examination for English and offers results and discussion for the first research question along with its sub-questions. The second section discusses the results concerning the relationship between the digital Matriculation Examination and current upper secondary English education. The third section summarizes and brings together all of the results, making connections to earlier research. As all of the teachers were women, I will use the female pronouns to refer to teachers and the male pronouns to refer to students.

4.1 Results about the digitalized Matriculation Examination

The results concerning the digitalized Matriculation Examination are presented in four subsections: 4.1.1 looks at the attitudes of the participants towards the reform, while 4.1.2 focuses on elements of the learning environment used for completing the example exercises. 4.1.3 considers the students’ experience on bringing out their language skills, and, finally, 4.1.4 analyzes the participants’ views on each example exercise in turn.

4.1.1 Attitudes towards the reform

The questionnaire included several Likert statements which aimed at finding out information about the participants’ attitudes towards the reform. In the below discussion, the attitudes that were revealed are discussed and juxtaposed with relevant background information, such as gender, evaluation of own English skills as well as IT skills.

First of all, when it came to evaluating the importance of digitalizing the Matriculation Examination, the students were hesitant. On the Likert scale ranging from 1 (completely disagree) to 5 (completely agree), the average was 2.9, with male students at a mean of 3.0; somewhat higher than that of female students (2.7). As Figure 1 below demonstrates, the answers were distributed quite broadly, with a third of the students not taking stance. This indicates that many students have not yet made up their minds about the importance of the digitalization, perhaps since its implications have not yet profoundly affected the participants’ upper secondary school studies and do not affect their own Matriculation Examination.
Figure 1 The importance of the reform

The average was understandably lower, 2.6, when the participants were asked about their willingness to complete their own Matriculation Examination in the digitalized format if they had the chance (see Figure 2 below). A little more than half of the students disagreed. The Matriculation Examination of most of the participating students will take place next year and they have already been prepared for it throughout their upper secondary studies. In this light, it is encouraging that as much as a fifth of the students would rather choose this format to complete their Matriculation Examination on the basis of their experience of the example exercises.

Figure 2 Willingness to complete own Matriculation Examination electronically
Female students were found less enthusiastic about bringing the digitalization into their own Matriculation Examination. In contrast, other background information, such as the student’s own estimates of his English knowledge or IT skills, were not found to correlate with the responses, which is in contrast with previous research by Hurme, Nummenmaa and Lehtinen (2013). The difference between the two studies is that in this study, the students could reflect on actual experience on the example exercises. This may have reassured some students with poorer IT capabilities of the fact that their IT skills will not hinder their success in the Matriculation Examination. Students who viewed their English or IT skills as poor were therefore just as self-assured, or bewildered, in the face of this reform as were those with strong confidence in their own skills. Figures 3 and 4 show the effect of gender and IT skills.

Figure 3 Willingness to complete own Matriculation Examination electronically, according to gender
When it came to the importance of measuring oral skills in the Matriculation Examination, the students were considerably less divided. With the average value at 3.8, 67% of the students agreed that measuring oral skills was important, a fifth could not answer and the remaining 12% felt that measuring oral competency was not important in the Matriculation Examination. Figure 5 shows the students’ answers on the Likert scale.

Figure 4 Willingness to complete own Matriculation Examination electronically, according to own estimate of English skills

Figure 5 The importance of testing oral skills in the Matriculation Examination.
A clear majority of the students therefore consider the Matriculation Examination as it is insufficient in measuring the students’ language capability. Although the participants remain hesitant about the benefits about the digitalization, they appreciate the measurement of oral skills brought along by it.

As illustrated by Figure 6 below, the oral dimension was considered an important addition even if the student was opposed to digitalization. The highest means for the importance of oral skills measurement are of students who took heavy stance on the importance of digitalization at both ends. One student who was opposed to the digitalization but supported the measurement of oral skills left the following comment in the free comment section of the questionnaire:

Q1 “I think the Matriculation Examination for English should not be digitalized! There are too many risks related to technology, which may ruin the whole test performance. An oral part could still be added to the test.”

This student saw the risk of technical problems interfering with test performance in the digitalized Matriculation Examination as too great, but still supported the addition of oral skills in the exam, which is acknowledged as a positive consequence of the reform.

Figure 6 Importance of testing oral skills in the Matriculation Examination, mean / Importance of digitalizing the Matriculation Examination
Regardless of what they thought about measuring oral skills in general, almost half of the students (47%) agreed to being happy about not having an oral dimension to their own Matriculation Examination. Students saw measuring oral skills as important on a general level, but would generally be reluctant to make it part of their Matriculation Examination. As the students called for more oral skills practice in upper secondary English teaching (see section 4.2.2), it is likely that most of the students feel unequipped to complete oral tasks in the Matriculation Examination and thus feel that this would lower their grade. Nevertheless, with 24% opting out of answering, 29% of the students disagreed to the statement, not viewing the lack of an oral dimension to their own test as positive. These students may feel that their test performance could be enhanced if they had the opportunity to show their oral skills.

Even though female students showed less enthusiasm for the digitalization overall, this inclination disappeared when it came to testing oral skills. Male and female students showed here almost equal results, with male students being slightly happier about not having to complete the oral part of the Matriculation Examination themselves.
It can be concluded that students have fairly mixed attitudes towards the
digitalization of the Matriculation Examination. As this was, for most, the first
concrete example of what the new Matriculation Examination could look like, it is
easily understood that opinions are yet to be formed. A clear majority of the students,
however, recognize the importance of measuring oral skills.

4.1.2 The learning environment

The ViLLE learning environment was perceived as a straightforward and
user-friendly platform for the example exercises. After approximately 5 minutes of
instruction, the students completed the exercises independently and presented few
questions on how to use the environment. With the mean at 3.8, 74% of the students
felt that it was easy to learn to use the software, even though 68% considered it
different from the programs or applications that they typically use. (See Appendix 8
for figures concerning these questions.)

In the questionnaire, the students had very few comments on learning to use
the software, which implies that students did not experience a lot of trouble learning
to use it. One student, however, expressed his insecurity concerning the use of the
software:

Q2 “When I opened the exercises, I noticed that regardless of the instructions, I wasn’t
sure what to do.”

This comment may, however, reveal more about the novelty of the experience for the
student than of the user-friendliness of the software.

Answers to the questions concerning the learning environment did not
correlate with students’ estimates of their own IT skills. Figure 8 below shows that
the highest mean for the statement “Learning to use the software was easy” was
equal of students who had estimated their IT skills as poor and those who saw them
as very good. If this is the case with software being introduced to students in the test
situation, it can be expected that using the Matriculation Examination software, the
use of which can and will be practiced beforehand, will not pose a problem for the
participants in the actual Matriculation Examination situation.
Learning to use the software, according to own estimate of IT skills

Figure 8

Given that the mean was relatively high for this question, the fifth that did not trust the software made their voice remarkably clear in the comment section. The
following comments were left in the free comment section or as answers to the question “What was bad about the example exercises?”

Q3 “The graphics were weird somehow childish to some extent”
Q4 “The website didn’t seem official”
Q5 “Unclearness and uncertainty about whether the answers are going where they should go.”

As ViLLE was designed as a learning tool for all ages, its graphics are according; they do not conform to the expected rigid black and white form of the paper-based Matriculation Examination. Rather, ViLLE strives to appeal to diverse student material and act as a tool, first of all, for learning (Laakso, Kaila and Rajala, 2014). The colorful graphics compared of ViLLE are markedly different from the paper-based Matriculation Examination, which was noticed by the student in Q3 and perhaps Q4, too. The Matriculation Examination software, however, as a system created specifically for this purpose, will surely be designed on different grounds.

Even though most students are not aware of the exact path that their answer papers will take after leaving their hands in the Matriculation Examination, either the software or the electronic format in itself seemed to cause additional unease about the security of his answers to the student quoted in Q5. The quotations underline the importance of suitable graphics in making the students trust the test software.

In their suggestions for improvement, students also hoped for a clearer, more user-friendly and more secure test.

Q6 “Making the system simpler and clearer is important.”
Q7 “the site could be made clearer also with regards to its looks, and it could be made to feel more secure”
Q8 “There is still a need for development in the system and its graphical user interface (gui).”

In addition to these general suggestions, the students came up with several specific suggestions to improve the user-friendliness of the test. One such example was a button for marking exercises as undone or demanding further attention.

Q9 “The text exercises could have a specific button you could press every time you’re unsure of some answer, so that when you’re checking, you would remember where you need to pay more attention.”

Even though a note like this could be typed into a word processor during the exam or scribbled down on paper, it is evident that it would be easier and handier for the student if the test software had an inbuilt way to do this.
A related quality highlighting the importance of checking that was brought up in several responses was being able to edit the answers until the student has finished the whole exam and is ready to submit all the answers. In addition, the students hoped for a fool-proof way to ensure that the test-taker cannot accidentally submit a wrong answer.

*Q10* “The “end test” button should be moved somewhere where you can’t click it accidentally so easily. -- You should be able to edit your answers until you have sent the whole exam and logged out of the service.”

*Q11* ”The "end test" button could be at the very bottom of the main menu and it could still ask for verification that the student wants to end the test. An x would be needed in the corner of multiple-choice questions, so that the windows could be closed without having to choose an answer.”

*Q12* “My attention was also caught by the time that kept running in the left-hand corner of the screen. Doesn’t a specific time frame just cause stress to the student, when you can’t decide yourself what to do first etc?”

The freedom to move around the exam exercises throughout the whole test time preferred by the students is somewhat in conflict with the current plan of the Matriculation Examination Board, according to which the language test will begin with the listening/oral part and then proceed in sections (reading comprehension, grammar, written production) which must be completed and submitted in sequence (Digabi, 2014). After submitting the answers to one section, the student can choose which section to complete next but cannot go back to edit the answers of previous sections. This will make the test more scheduled than the paper-based one and limit the student’s freedom in the test situation, which, according to Q12 above, might “cause stress to the student”.

On the other hand, one student brought out his preference for an automatically proceeding test:

*Q13* “The software could be simpler and it could proceed like “on its own” so that the student doesn’t have to keep clicking new places, because it makes you confused”

This vision, in turn, is supported by the current plan, since the system will direct the student through the exam along an automatically guided path if he does not choose which section to complete first within a given time frame (Digabi, 2014). Time frames therefore get both support and criticism from students, although the flexibility of the software is supported by more students.

Another relevant viewpoint was brought by a student with dyslexia, hoping for better support for students with special needs:

*Q14* “Bigger and thicker fonts, for someone with dyslexia it’s extremely difficult to follow the text without color foils and without highlightings or underlinings.”
These issues had not been considered in the research design of this study but they are, of course, crucial to take into account when dealing with the large diversity of test-takers in the Matriculation Examination.

To summarize, the answers show that the ViLLE learning tool is a user-friendly and reliable learning environment appealing to users of all levels of IT capacity. Learning how to use it does not require experience from using similar software. Whereas it may be nice for versatile learning software to include bright colors and graphics that appeal diversely to all age groups, many participants expect the Matriculation Examination software to be more “official” than ViLLE when it comes to graphics. Graphics affect the perceived reliability of the software which adds to the students’ feeling of security in the test situation. Students hope for a clear, straight-forward test software which would proceed logically but also offer the possibility to complete the test in their own preferred order. The ideal test software would also protect the student from unintentional wrong answers, both when it comes to technical errors (clicking the wrong button leads to a confirmation message) or linguistic insecurity (the student can mark the answer as uncertain and come back to it at the end).

Based on the discussion in 4.1.1 and 4.1.2, we can conclude, first of all, that gender plays a role in the CBLT experience: female students were less inclined in all respects to support the digitalization or to view the software as user-friendly, easy or familiar. This is demonstrated in Figure 18. In contrast, other factors, such as level of English skills or level of IT know-how were not found to play a significant role in students’ experience of the example exercises. This indicates that the example exercises are not biased against poorer students, neither with regards to subject knowledge nor to IT skills.
Figure 10 Summary: familiarity or degree of enthusiasm towards the reform and test software, according to gender

4.1.3 Measuring and bringing out language skills

Another interest of this study was to find out how students feel their language skills are measured and brought out in the new exam format. It was found that the majority of the students (59%) felt that the example exercises measured their language skills in a diverse way, the average value being 3.5. In their answers to the open questions, however, many participants remarked on the dominance of reading comprehension, which made for 50% of the example exercises, and some students mentioned that they would have liked to try examples of listening comprehension, too. Due to the limited time available for the completion of the exercises, the task types were restricted to four and they could not be made as versatile as the Matriculation Examination for English is (see 4.3 for a more detailed account on designing the example exercises).

Still, altogether ten students mentioned the word “versatility” in their answers to the open question “What was good about the example exercises?” Judging by the students’ comments (discussed in more detail in 4.1.5), the added versatility of the example exercises compared to traditional Matriculation Examination exercises derived from the blog exercise (example exercise 1) and the oral task.
As for the perceived authenticity of the example exercises, the majority of the students (62%) agreed that the exercises represented situations which they could face outside the test situation. The differences between the exercises might affect these results, since some exercises were likely to be perceived as more authentic presentations of real-life language use situations than others. Still, the exercises overall were considered by the majority to test language skills in an authentic way.
The participants were also asked to assess whether the digitalized Matriculation Examination seemed to them similar to the paper-based exam. As much as 71% disagreed, with the average value at 2.3 for the statement. This indicates that the students regard the reform as a significant change to the English Matriculation Examination. Figure 13 illustrates this finding.

![Graph](image)

**Figure 13 Similarity between the computer-based Matriculation Examination and the paper-based one**

In their answers to the open questions, three students mentioned the similarity between the digitalized and paper version.

Q15  “Some were similar to the tasks in the paper version but did not include flipping through papers etc.”

Q16  “The structure of the exercises was nearly identical with that of the paper-based Matriculation Examination.”

Q17  “They [the exercises] were similar to the normal exam.”

These quotations are all taken from the students’ answers to the question: “What was good about the example exercises?”, which means that the familiar aspects of the exercises were considered positive. Still, the comments acknowledge some differences between the formats. In Q15, the participant mentions the enhancement of not having to flip through papers. Q16 considers the structure of the exercises familiar, but does not elaborate on the contents here. In addition, one student mentioned the “new way” (orig. Finnish “uudenlainen tapa”) as a positive side to
the exercises. The students then seem to appreciate both a certain degree of familiarity and some new enhancements of the digital test format. Still, it is evident that the digitalization affects the students’ experience of the exam a great deal.

Turning our attention to the students’ experiences of their own test performance, students generally felt they performed similarly or slightly worse when completing the test on the computer than on paper. When asked how well the students felt they could bring out their language skills on the computer when compared to the paper exam, the average on a 5-point scale was 2.8, with 40% of the students evaluating their performance as equal in both forms. Figure 14 presents this finding.

![Figure 14 Bringing out language skills, CBLT vs PBLT](image)

Some students were so aware of the test format that they felt that it overshadowed their language skills. One student phrased his feelings in the following way:

**Q18** “They were difficult because there were so many links and tabs. It made me feel I know a lot less than I had thought.”

This student clearly linked his inability to immediately master the use of the test software with his language ability and felt discouraged in both skills by this experience.
The majority of the participants, 62% felt that they learned as much during the completion of digitalized Matriculation Examination exercises as during paper exams. 23%, however, felt that they were learning less than on paper, whereas 15% experienced the contrary, setting the average value at 2.8. The high frequency of the middle answer may also be due to the question. This was perhaps one of the hardest things to assess in the questionnaire.\footnote{Moreover, the question was somewhat awkwardly formulated: the resulting sentence would have been “In the electronic exam situation, I feel like I am learning while completing the exam worse than on paper”, whereas the words “more” or “less” would perhaps have fit the sentence better. This may also have guided some students towards the center of the spectrum.}

![Figure 15 Learning while completing the exam](image.png)

One of the participants who agreed to this statement left the following comment:

\textit{Q19}  “Interesting texts, for which well-suited questions, joy of learning”

Though this learner does not mention the format in which the texts and questions are, it is clear that the digital test offered him at least as good a platform to experience this joy as a PBLT would have.
When it came to concentration, the majority of the students felt that their concentration was worsened by the digital format. With the average value at 2.4, altogether 53% assessed their concentration as worse in the digitalized format. In the test situation, the overall difficulty in concentrating could be explained by the researcher’s further instructions to individual students who asked for them in the test situation. Also strict time limits were reported to stress the student in the test situation of this study, which may affect the students’ overall performance and working efficiency. Then again, some students finished early and as access to internet was not restricted to the ViILLE server, were found surfing online during the test. More research is therefore needed to support this finding.

![Figure 16 Ability to concentrate, CBLT vs PBLT](image)

Figure 16 Ability to concentrate, CBLT vs PBLT

When asked how they felt they performed in an electronic test situation in comparison with a traditional test situation, women evaluated their performance as worse than men did theirs. As when comparing attitudes concerning the reform and the test software, women’s experience was inferior to men’s in this regard as well. This is demonstrated by Table 2.
In an electronic exam situation,

<table>
<thead>
<tr>
<th>Gender</th>
<th>Man</th>
<th></th>
<th>Woman</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
<td>Std.</td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>... I can bring out my language skills</td>
<td>2.9</td>
<td>.9</td>
<td>2.7</td>
<td>1.1</td>
<td>2.8</td>
<td>1.0</td>
</tr>
<tr>
<td>... I can work carefully</td>
<td>2.7</td>
<td>.8</td>
<td>2.4</td>
<td>1.0</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>... I can concentrate on my work</td>
<td>2.5</td>
<td>.9</td>
<td>2.2</td>
<td>.8</td>
<td>2.4</td>
<td>.9</td>
</tr>
<tr>
<td>... I can make efficient use of the time given to complete the exam</td>
<td>3.0</td>
<td>.7</td>
<td>2.7</td>
<td>1.1</td>
<td>2.9</td>
<td>.9</td>
</tr>
<tr>
<td>... I feel like I am learning while completing the exam</td>
<td>3.0</td>
<td>.7</td>
<td>2.6</td>
<td>.9</td>
<td>2.9</td>
<td>.8</td>
</tr>
</tbody>
</table>

Table 2 Experiences on working efficiency in an electronic exam situation, according to gender

These questions were considered especially interesting to look at in relation to the student’s IT skills. It was found that in questionnaire items 15, “I can work carefully”, and 17, “I can make efficient use of the time given to complete the exam”, the answers correlated with the students’ estimate of their IT skills. The poorer the student’s IT skills, the poorer he felt his capabilities were in these respects.

Figure 17 Experiences on working efficiency in an electronic exam situation, according to IT skills
As is evident in Figure 17, this correlation was especially strong in the question about working carefully: with 5 meaning “significantly better than on paper” and 1 “significantly worse than on paper”, the mean for students with who estimated their IT skills as “poor” was 1.9. For those who estimated their IT skills as excellent, the corresponding mean was 3.0. For making use of the given test time, the gap between the same student groups is almost equal: 2.3 for those with poor IT skills, 3.3 for those with excellent ones. The students who considered their IT skills as excellent surprisingly assessed their capability to concentrate on their work as the lowest, even though this correlation was not as strong as the two others.

The results concerning the students’ assessment of their own performance in the digital exam indicate that students have yet to learn to reach their best performance in digital test surroundings. However, since the data gathering for this study was not ideal for assessing test performance in that the tasks were completed as part of normal schoolwork, more research is needed on the students’ performance in test-like conditions.

4.1.4 The example exercises according to task type

The following discussions will center on each example exercise in turn, all to be found in the Appendices. The data for these discussions derives from the open-ended questions of the survey: the students were asked to describe each exercise using one to three words, and in the end they were asked about the good and bad sides of the example exercises.

**Example exercise 1: Reading comprehension (multiple choice)**

The multiple-choice reading comprehension exercise received, alongside the oral skills exercise, the most feedback from the students. This exercise consisted of six multiple-choice questions which dealt with the contents of a news site created on a blog platform. This news site included four news articles, and the questions did not always make clear which article the answer was to be found in. Sometimes answering a question required consulting multiple articles.

---

10 Note: the illogical frequency of the words "green", "slimy" and "likeable" is explained by the fact that they were used in the questionnaire as example words in the instruction. Students probably used these words to be humorous or to avoid having to come up with an answer of their own for the question. When it comes to “likeable”, orig. Finnish “sympaattinen” it is difficult to know whether the word was used sincerely or not.
Authenticity was emphasized in the students’ descriptions of this exercise. Words like “modern” and “topical” occurred frequently in the students’ descriptions. Still, the most repeated words were “interesting”, “confusing” and “good”. Picture 1 illustrates the words used to describe the exercise. The size of the word in the graph corresponds with the frequency of the word in the data: the bigger the word, the higher its frequency (see Appendix 9 for full-page graphs of Pictures 1, 2 and 3).

Picture 1 Students’ comments on example exercise 1: Multiple-choice reading comprehension

As the most prominent word in the graph implies, the students found the exercise very interesting. This can apply to both the topics and the format of the exercise, both of which were distinguishably different from the traditional Matriculation Examination reading comprehension exercises.

The topics were considered interesting as they dealt with more current issues than what normally makes the Matriculation Examination. The following quotations come from the students’ answers to the question “What was good about the example exercises?”

Q20 “They were about topical issues and topics that interest people of our age.”

Q21 “The texts were interesting, and not too dry expository texts. Their topicality, too, made them easier to understand”

The reason behind using older texts for the Matriculation Examination might be reducing the effect of the student’s general knowledge of topical issues on the performance and minimizing the number of students working with a familiar text. However, according to these comments, dealing with texts concerning current issues increases student motivation and sense of learning in the exam situation.

Besides the fact that the topics were current, the format also makes a difference. A number of students appreciated the fact that real news and articles were
utilized in the exercise, even though that has long been the case even with the paper version of the exam.

*Q22*  “They were more about the real world for example news were used in the reading comprehension exercises”

Seeing the news in an everyday online surrounding made the students feel that the material is more authentic, more “about the real world”. When asked to describe the exercise in one to three adjectives, one student described it as “depicting an everyday situation” and three students used the word “practical” to describe its hands-on nature. In addition, one student stated:

*Q23*  “Reading comprehension exercises were good, because they measured your skills in reality.”

Even though the words “in reality”, or “really” (translation for the Finnish word *oikeasti*) are not elaborated on, there is a clear juxtaposition of the digitalized Matriculation Examination, which “really” measures skills, and something else, perhaps the traditional format of the Matriculation Examination, which does so somehow inadequately. The fact that students observe the increased authenticity of the example exercises supports the view that authenticity will be enhanced in the digitalized Matriculation Examination (see section 2.2.5).

Students also noticed the fact that this exercise measured modern-day language skills in a new way. For example, media literacy was tested in one question. To be able to tell that one multiple-choice alternative was wrong, the students had to check each article to see whether their source was made transparent. In addition, it was not made clear in the instructions where the answers to the questions could be found, which meant that the student had a large material wherein he had to look for the relevant information. Dealing with a large base of source material is a present-day language skill representing real life better than working on a single, short text or paragraph. This kind of measurement was received with both praise and bafflement:

*Q24*  “In the online news you had to use your brain a little when they didn’t tell you directly where to find the answer. They were however quite limited in range, but it’s likely that the future Matriculation Exam will include more versatile questions.”

*Q25*  “It was a little difficult to follow which question dealt with which article (in the first task).”

Both students acknowledge that new skills were being measured: one felt he possessed the skill and found it challenging in a new, exciting way, while the other felt unequipped to deal with the abundance of the source material. The latter feeling is consistent with Lakkola and Ilomäki’s (2013) claim that modern-day media
literacy skills have not yet reached the language classroom (also supported by Hautamäki et al., 2012; Kiili, 2012; Pajarre, 2012).

What was clearly a defect in the blog exercise was the fact that the questions and the source material had to be viewed in separate windows or tabs on the computer. Some students did not see this as a problem as they adjusted both windows to fit their screen simultaneously. The design, though chosen because it was not possible to integrate the type of versatile news platform to the test software itself, can be seen as representing a real-life situation where multiple information sources are at use simultaneously. Still, the system was fairly inconvenient, which was remarked upon by a large number of students. The frequent instances of the words “confusing”, “unclear” and “tricky” in the students’ descriptions are doubtless partly due to the awkwardness of the multiple windows. One student also aptly described it as “fragmented”.

Q26 “I think the first exercise had too many tabs, which may confuse the student.”

Q27 “The menu in the multiple-choice task for reading comprehension was weird and there were a bit too many sites to open”

These answers support the findings presented in 4.1.2 that students wish for a straight-forward, user-friendly software which would minimize unnecessary browsing and clicking.

One student found pictures disturbing, although including pictures and audiovisual material is regarded as one of the distinguished richnesses of the digitalized format (see e.g. Choi, Kim and Boo, 2003).

Q28 “In the reading comprehension the pictures were disturbing concentration and they could be left out.”

The quotation highlights the individuality of the test experience. Nevertheless, it may also be a symptom of shunning something that is new. Any kind of multimedia is seldom present in paper-based tests, which means that including these elements in abundance may well distract or seem useless to some students in the beginning. However, since pictures and audiovisual material make for a central part of both present-day language use and language learning, there is no reason for it to be absent in language testing situations. Picture 1 also demonstrates the effect of individual experience: the antonyms “clear” and “unclear” are of equal size in the graph, which means that they were mentioned equally often (both by altogether seven students). Whereas some students experienced the exercise as unclear, others perceived it as
straightforward. The words “difficult” and “easy” also both occurred rather frequently, with the word difficult often occurring in a longer description, such as “difficult to use” or “difficult to understand”.

**Example exercise 2: Reading comprehension (open questions)**

The second reading comprehension exercise featured a short article, a picture and five open-ended questions in Finnish. This is a traditional task type in the Matriculation Examination which has been transferred into the electronic format with few additions. Picture 2 below illustrates the descriptions of students.

![Word cloud for students' comments](image)

*Picture 2 Students’ comments on example exercise 2: Open questions*

The exercise format was familiar to the students, and the task overall was described in pleasant terms. As can be seen in Picture 2, the exercise was, above all, considered “easy” and “good”. The words “clear” and “nice” as well as “ordinary”, “difficult” and “boring” also stand out.

Even though the difficulty level of the language content of the exercises is not relevant for this study, other meanings of the word “easy” became significant in the students’ comments. As this was the only task demanding written output, the following comments of students refer mostly to this task:

*Q29*  “Answering was easy and fast”
Q30 “You could write on the computer. If you made mistakes it was easy to just press delete.”

The format of answer output was therefore considered easy and user-friendly. The word “easy” was also brought out in the following comment comparing the two reading comprehension exercises (answer to question: “How do you think the digitalized Matriculation Examination for English could be developed?”)

Q31 “In the reading comprehension the articles would be displayed in an easier way (like in exercise 2)”

The word “difficult”, however, seemed to refer mostly to the difficulty level of the vocabulary and questions, since no one used it in their further comments of the exercise. In contrast, the difficulty of the vocabulary and questions in all of the exercises was commented on by several students. The task format and display were therefore considered easy and user-friendly.

In addition to and perhaps as a consequence of this perceived easiness, students also used the words “good”, “clear” and “nice” to describe the exercise. One student also mentioned this task as a whole as a good side of the example exercises. In total, students liked the exercise and described it in positive terms.

However, the exercise aroused little more than lukewarm feelings in the students. Prominent were also words like “ordinary” and “simple”, which can both be considered both positive and negative, as well as the more unfavorable “boring”. Comments that mention this exercise specifically remain considerably scarcer than those which mention the first exercise, for example. Since most students completed the exercises in sequence, this exercise may well have been somewhat boring after the more radically different first reading comprehension exercise.

Despite the fact that the exercise was not considered overwhelmingly interesting and innovative, it is clear that with its familiarity, easy display and facilitated answer construct, this task type was perceived as an appropriate and student-friendly task type in the digitalized Matriculation Examination. Students also saw this task type as the readiest task format in itself, since it received no suggestions for improvement.

Example exercise 3: Grammar and vocabulary

The third exercise was a short grammar and vocabulary task. It consisted of a cloze text with ten gaps which the student had to fill by choosing the right answer
from a drop-down menu. This was considered a feasible way to complete a task which is familiar from the paper-based exam.

![Word Cloud]

*Picture 3 Students' comments on example exercise 3: Grammar and vocabulary*

Much like the reading comprehension exercise with open questions, this familiar task type transferred to the electronic format was considered “easy” and “good” above all. Also the words “short”, “ok”, “challenging” and “difficult” are conspicuous in Picture 3.

A clear asset of bringing the familiar exercise into the electronic format was the fact that the selected answer showed up instantly in the gap left for it and the text could then be read as a whole. The below comments are answers to question “What was good about the example exercises?”

- **Q32** “answering did not take time, it was easy to just click and it was nice when in the cloze exercise in the grammar section the completed text could be read in its entirety”

- **Q33** “it was easy to change the choices in the cloze exercise, clarity in the cloze exercise, when you did not have to check to see whether the ball is at the right number”

As this is a traditional task type in the Matriculation Examination and English textbooks in general, many students made a comparison between the digital and the paper-based version of the task. In all the comparisons, the digitalized version was considered a more feasible way to complete the task. The student could work within the text itself and the risk of unintended errors was diminished, reducing the stress
related to it. As brought out by Q33, this reduces stress and insecurity in the checking process.

Due to the limited time available for the completion of the exercises, this part was purposefully made short. It was assumed that the practicality of the exercise type could be assessed even by means of a shorter task. Together with the fact that this was the final exercise completed in the students’ own pace, this explains the frequency of words like “short” (often occurring with “too”) and “unfinished”. The words “ordinary”, “normal”, “average” and “ok” again demonstrate the students’ familiarity with the exercise. The only criticism this exercise received had to do with the language content or the length of the text: the format was considered ready in itself. On the whole, then, students were at ease with this recurring exercise format, appreciated the facilitated answer mechanism and felt that it suited its purpose well.

Example exercise 4: Oral skills

The oral exercise was a new task type which the students could not relate to their experience of the paper-based Matriculation Examination exercises. The novelty of the experience is shown in the words of Picture 4, which demonstrates a wide range of adjectives ranging from “fun” to “unpleasant”, “useless” to “important”, etc. Many students took stance on the realization of the exercise: e.g. “badly” or “lously realized” were found among the comments.
The exercise obviously inspired plenty of comments, opinions and suggestions for improvement. Only 15 students did not present voluntary additional comments on the oral exercise in the open-ended questions. The remaining 78 made remarks on 1) the test situation, 2) equipment, 3) feelings that the exercise aroused, 4) preparation levels and 5) improvement proposals.

First of all, the implementation of the oral exercise in the test situation received a lot of criticism from the students. A third of the students mentioned some aspect of the oral exercise in their answer to question “What was bad about the Matriculation Examination example exercises?” What came up the most was the inconvenient test situation: students could not fully concentrate on their own performance, because the noise coming from others completing the exercise at the same time was disturbing them. Two students described the exercise as “lously realized”, one as “badly organized” and several as “unpleasant”, “tricky” and “unclear”. Even though all students started the exercise at the same time, some students experienced technical problems or started to rewind the audio instructions on their own, which led to students completing the exercise in slightly different paces. This resulted in some students talking whilst others were listening to instructions, which made it difficult for these students to concentrate on the instructions as the headphones were by no means soundproof. In their answers to the question “How do you think the digitalized Matriculation Examination should be developed?” students often mentioned these issues:

Q34 “the listening exercise, because it was really disturbing when others talked at the same time and you couldn’t hear the questions properly”

Other students were indeed the factor which most weakened the students’ ability to concentrate. Some students suggested the possibility to complete the oral part in turns or in separate rooms as a remedy for this, a vision unattainable in its impracticality.

Q35 “The oral exam should be arranged either in separate rooms or using noise-cancelling headphones.”

Even though the first suggestion in Q35 cannot be implemented, the quotation brings us to the second point brought out by many students, namely equipment and exam space arrangements. Like in Q35, several students hoped for headphones of better quality and soundproofing. Students also hoped for test software where the
recording facility would be integrated, because using a separate recording program for the purpose was regarded as tedious.

Q36 “It was painful listening to the instructions for the oral exercise from crappy earphones”

Q37 “I do not think the oral part was good. Headphones should cut off noise notably better in order for it to work.”

Q38 “In addition, the program could have its own recording facility and noise-canceling headphones for the oral part. (so that cheating wouldn’t be as easy)”

For an unfortunately high number of 22 students, malfunctioning technology led to being unable to successfully complete the oral exercise. This explains the frequency of the word “malfunctioning” in Picture 4. Nine students reported in their answers that their answers were lost due to technical problems. Six students suffered from a test arrangement mistake, where the IT classroom of this group had two types of computers functioning with different operating systems, and the recording facility was only tested on one of the computers, operating on the newer system. After the completion of the oral exercise, it turned out that the older operating system was using the internal microphone to record instead of the external one. In consequence, the students’ own recordings were lost and the computer only recorded the instruction audio. In addition to this technical problem, one student reported that he was unable to save his recording because his own username did not have enough space to save the recording. As a USB stick was at hand for difficulties like this, the instructions to ask for help in case of difficulty were not clear enough in this test situation.

In addition, due to lacking instructions or badly functioning microphones, some students held the microphone so far from their mouth that their responses were almost inaudible. Some students, either by accident or on purpose, submitted blank files. However, only two students suffering from broken headphones could not complete the exercise at all. Others still had the chance to do the task and reflect on this experience for the questionnaire, even though some were demotivated due to the problems they had faced. Out of 94 participants, 63 students’ submissions were audible and enabled assessment, whereas 22 students’ files were unfit for assessment. They had either submitted blank files or reported technical problems which inhibited them from submitting their file. The remaining 9 students did not submit a file for unknown reasons. The oral exercise, therefore, was completed at a success rate of 67%. Many of the failures took place during the first two test sessions, whereas the
final test session enjoyed a success rate of 100%. Part of this is doubtless due to improvements in clear instructions from the researcher as well as the increased ability to help students deal with unexpected problems occurring in the test situation.

These setbacks demonstrate the problems that a language teacher may encounter when utilizing ICT in her teaching, and they gain more importance when ICT is used for testing, e.g. course exams. Although these very problems are unlikely to take place in the Matriculation Examination, it is certain that some technical problems may affect the actual exam situation as well. Identifying potential problems and assigning and equipping appropriate people to deal with them are surely challenges demanding, and currently receiving, a lot of attention from the Matriculation Examination Board.

Thirdly, with reference to feelings provoked by the oral exam, students reported some amount of oral test anxiety and even anger. Anxiety was perceivable in some students in the test situation. In the questionnaire it showed as more colorful and even offensive language being used when talking about the oral exercise. Students characterized the exercise e.g. as “upsetting”, “nerve-racking” and “causing panic”. These word choices show the strong emotional aspect of the exercise. Several students felt that it was embarrassing to complete the exercise if they were lagging behind others in time. Anxious students expressed their discomfort in different ways in the free comment section:

Q39  “the oral exercise was somehow distressing”

Q40  “The oral part was not realized in a very practical way. The noise in the classroom made me forget for example how I had started the sentence or what I was about to say and also made hearing difficult. In the actual Matriculation Examination it should be better realized. It should also be considered that the student might not succeed as well in the oral part if they have a shy temper and do not enjoy performing. When you are anxious in the test situation, blood hums in your ears and your mouth is dry, which makes talking more difficult.”

Q41  “You can’t do something like this if the earphones are that bad, go and buy a [f******] 10 E headset at Prisma and you’ll hear better than out of those cancer gimmicks”

While the student quoted in Q39 is unable to further determine his feelings of distress, Q40 analyzes closely his feelings and experience, identifying the causes on multiple levels. The student cited in Q41, in turn, clearly expresses the frustration he experienced in the test situation. This kind of language was not found in the descriptions of any other exercise. It is apparent that this new task type aroused the strongest feelings in the students.
Still, there were a lot of students who found the oral test especially pleasant – indeed, “fun” was the most frequently used word to describe the oral exercise, with twelve occurrences. The words “nice”, “good” and “great” also had high frequencies.

Q42 “The oral test was nice :)”  
Q43 “The oral test was good in that you got to show your speaking skills.”  
Q44 “I think testing oral language capacity was EXTREMELY GOOD! Indeed, that’s exactly why we learn languages so we could use it in real life as well.”

The diversity of feelings aroused is closely linked to the preparation level of the students, which was the fourth recurring theme in the students’ answers. Part of the fun and part of the anxiety was most likely due to the fact that the oral exercise was a strange task type for the students. The situation was surely new for the students and the fact that they had no time to rehearse or formulate their answers beforehand made it even more challenging. One student acknowledged the effect of the novelty factor:

Q45 “unaccustomedness to completing an oral test also evoked confusion but had the task type been practiced beforehand, it would be less confusing”

It is evident that the oral part, however implemented, will include additional stress factors and concentration issues, especially in the beginning of its operation. Some of these issues were identified by students:

Q46 “In the oral part you can’t correct yourself if you get it wrong.”  
Q47 “... in the oral part the result can be remarkably worsened by anxiety in the test situation, which may cause for example stuttering.”

Nevertheless, the student must have a guaranteed chance to hear or view the instructions and not be excessively disturbed by other students carrying out their own performance. Ensuring this chance can be done through careful planning of the exercise type, test situation and exam space arrangement and through correct tools. The exercise type used for this oral exercise was perhaps distressing in that it gave the student no time to prepare his answer beforehand. Even though part of the oral exercise can test spontaneous reactions to instructions, the effect of stress factors identified in this study might be diminished if at least some of the tasks in the oral part would include time for the student to prepare his answer before carrying out the task.

Finally, while the students by no means saw the task type as ready, plenty of students showed clear motivation towards developing the important task type further.
by carefully articulating the deficiencies of the implementation of the oral exercise, even if they sometimes lacked concrete ideas on how to fix them.

Q48 “The oral part is however almost the most important thing to know about the English language and you need it the most. Its testing should be increased and developed a lot.”

Q49 “Organizing the oral exam was poor in the test situation. For sure, hearing others speak and the resulting cheating and disturbances will probably be minimized in the real exam situation, but in this test situation hearing others speaking was very disturbing to me. Often I couldn’t hear the instructions, and it was difficult to repeat example sentences, when I’d only heard half of them.”

Furthermore, there were students who suggested more versatile measurement of speaking skills:

Q50 “I think there should be more listening comprehension and discussion situations”

On the other hand, there were also students who questioned the validity, practicality and importance of the oral part. Three students described it as “useless”.

Q51 “I wonder if the oral part is sensible.”

Q52 “The oral exam was impractical and needs improvement. I myself feel that the oral exercise is unnecessary.”

As for the test situation and exam space arrangement, it would be beneficial for students to not face each other in the oral section, since this would ensure better concentration on their own work. Finally, the proper tools would enable the student to hear the instructions and cut out the noise from other students.

The aim of experimenting with the oral task for this study was to utilize tools which already exist in schools and test whether they, as such, could be of use in the completion of the Matriculation Examination. The results indicate that though these tools can be of considerable use in the process of preparing the student for the Matriculation Examination as part of classroom work, they are inadequate for the actual completion of the oral task of the Matriculation Examination. The introduction of the oral exercise will therefore require the procurement of new tools along with creating strict guidelines for their implementation in the Matriculation Examination in order to ensure that all students get an equal chance to complete the oral section. As the aim of this study was to provide a look into students’ views on the overall reform and its influences, it is but a scratch on the surface of a topic demanding plenty of more research specifically dealing with finding appropriate implementation models for the oral part, testing different tools for its realization and documenting and analyzing test-taker experiences from diverse user groups.
4.1.5 **Overall user experience**

On the whole, the experience of completing Matriculation Examination exercises on the computer seemed to appeal to the participants. Numerous students commented on the handiness and agreeability of working on the computer. Typing and correcting mistakes were the qualities appreciated the most.

Q53 “It was handy reading off the screen. Nicer to complete than the paper version.”
Q54 “rather handy and you don’t have to write by hand”

Correcting mistakes was rendered easier through the digitalized format. This was brought out in several comments.

Q55 “It’s faster to work on the computer and for example correcting mistakes is easier.”
Q56 “It was easy to kind of ‘erase’ if you answered wrong.”

A more general opposite view was brought out in the comments section by only one learner:

Q57 “It’s overall worse to do exercises on the computer.”

Although individual preference naturally plays a role in the experience, the widely-held opinion among students was that the completion of the exercises was rendered faster, handier and more modern by the new electronic format.

While some were opposed to the digitalization altogether, a small number of students saw the software and realization as outdated already.

Q58 “The realization was extremely old-fashioned but the idea was good”

Naturally, the version the students made use of had deficiencies which will not be present in the actual Matriculation Examination software (such as the aforementioned multiple tabs in the reading comprehension as well as using a separate recording facility for the oral skills part). Still, these comments bring our attention to the fact that software as well as hardware used in the Matriculation Examination is quickly outdated and hence demands constant updating in a different way than supplies for the paper version.

4.1.6 **Suggestions for improving the digital Matriculation Examination exercises**

When moving to improvement suggestions regarding the overall contents of the new exam, students had a myriad of concrete ideas. Out of Bachmann and Palmer’s (1996) test usefulness qualities, students especially appreciate authenticity
and interactivity in their comments. For example, some students interpreted the lack of a traditional listening comprehension section in the example exercises as the potential lack of it in the actual digitalized Matriculation Examination, and brought out their hopes of preserving this section and enhancing it with video input and interactivity:

Q59 “I think the test could include something that corresponds to the earlier listening comprehension section (though shorter), because I guess at some lectures you need to be able to listen closely to long speeches as well.”

Q60 “an oral communication exercise where the aim is to respond understandably to different situations which you could face on the street for example.”

Q61 “I think moving picture could also be included e.g. TED interviews and videos”

Quotations 59 and 60 markedly emphasize authenticity by making direct links between a situation to be faced in further studies, possibly abroad, as well as one that could occur in everyday life. Q61, in contrast, calls for more interactivity of format in the exam content.

For some students, the solution to problems experienced with the digitalized test was backing out of the reform:

Q62 “Only the answers on the computer”

Q63 “The Matriculation Examination should not be digitalized. Computers can crash any time, and it does not feel as natural as writing with your own little hands. On top of everything, the contents of the examination will change for the worse along with the digitalization.”

Behind these suggestions may also be a fear of the far-reaching impact of this reform, which was brought out in one comment:

Q64 “If all Matriculation Examinations are digitalized, people will lose the ability to write by hand.”

Many students, however, saw potential in the example exercises and expressed belief in the success of the reform or willingness to help develop the exam further:

Q65 “[it should be developed] towards a bit more familiar direction, but maintaining some features, such as the demand for heightened information analysis.”

Q66 “When students are familiarized with the digitalized Matriculation Examination and the exam tasks are developed a bit further and the system made better, the reform will surely be successful.”

All in all, presenting the students with an early version of the implementation of digitalized Matriculation Examination example exercises helped the students identify important deficiencies and make valuable suggestions for the future development of the reform.
4.2 Results about the relationship between foreign language education and the digitalized Matriculation Examination

A secondary research aim was to find out how students perceive the relationship between the English education they are currently receiving and the example exercises from the digitalized Matriculation Examination. The students were asked about their current English teaching, their views of good English teaching and what, if anything, should be done differently in English teaching once the Matriculation Examination is in the electronic format.

4.2.1 Results about current language education

Students seemed to be generally happy about the linguistic contents of their English education. The large majority (80 %) felt that the teaching they are receiving develops language skills which will be needed in future studies and working life, with the mean value at 3.9.

However, the use of ICT in English teaching was wavering. Only a little less than half of the students (47 %) felt that their English education develops the technical skills which are needed in language use situations, with almost a fourth disagreeing to the statement. When it came to assessing the efficiency at which technology is utilized in teaching, students were split: 41% agreed that its use was efficient, whereas 40 % thought the reverse was true. The mean values on the Likert scale for these statements were 3.3 and 3.0, respectively.

Another question which divided opinions was evaluating how well the English teaching they are receiving prepares the students for this type of exam. 38 % felt that it prepared them well, 26 % could not answer, and 36 % disagreed to the statement. As for learning material being used in English teaching in the upper secondary school, 49 % found that it makes use of exercises which are similar to the example exercises presented to the students in this study. 35 % opted out of answering and 15% disagreed, leaving the average at 3.4 on the Likert scale. Figure 18 shows the means for the students’ answers to the questions concerning their current English education.
The differences in student opinions may derive from differences between teaching groups, which varied greatly in terms of ICT utilization. Another factor which affects the results is the student’s own attitude to the example exercises. Some students were more inclined to look beyond the format of the exam and instead into the contents, which they saw as similar to the paper-based Matriculation Examination which they are currently being prepared for. Others, however, saw the change as the revolution of the entire exam, demanding an all new type of English education.

Figure 18 Current English education

Figure 19 What students consider important in upper secondary English Education
Figure 19 above shows what the students considered important in English teaching. The students gave high value to practicing the task types of the Matriculation Examination, imitating outside world situations and interaction with other speakers of English. When it came to online skills and computer use in English classes, the averages remained relatively high, with 42 % of the participants agreeing that it is important for the student himself to use the computer in class. Figure 19 illustrates this finding.

These answers show that in addition to valuing authenticity and interactiveness especially as test qualities (Bachmann and Palmer, 1996; see 4.1.4), students regard them as very important qualities in English teaching and learning, too.

**4.2.2 Results about required changes to English education**

Students were inquired about the changes they see as important in upper secondary English education in the future years through a series of Likert statements along with an open-ended question. It was found when reviewing the data that the open-ended question suited this purpose better, as the students got to phrase their own answers without being forced to take stance on a ready-made statement. Figure 29 still exhibits the students’ answers to some commonly predicted views of the future of English education which were chosen for the questionnaire.

![Figure 20 Developing upper secondary school English education](image-url)
Students were almost unanimous about the fact that the demands of the Matriculation Examination should affect the contents of upper secondary education. The attitudes concerning the relationship between the Matriculation Examination and upper secondary education have not been previously researched in Finland, but the question of how much of upper secondary school education should revolve around specifically preparing the student for the Matriculation Examination is often discussed among teachers. The students’ stance, however, was clear: 89% agreed that it was important to practice the task types of the Matriculation Examination.

When considering the following quotations in this light, it is clear that students’ English education as such does not offer chances to practice the digital task types:

Q67 “The test was otherwise good, but it felt a bit weird compared to the paper version.”
Q68 “For now, the paper version seems more natural, even though it’s a question of getting used to.”
Q69 “takes a little getting used to”

These students acknowledge the fact that familiarity to test tasks plays an important role.

However, a more critical view was also brought out by one student in the free comment section:

Q70 “We shouldn’t concentrate simply on studying the Matriculation Examination all the time but we should put more effort into learning to manage with your own English language if you really need to use it.”

This quotation emphasizes authenticity, a quality which constantly resurfaces in the students’ answers. The student underlines the importance of becoming comfortable with authentic language use situations and learning to deal with them. As argued in 2.2.5 and found in 4.1.4, the digitalization of the Matriculation Examination enables new kind of authenticity in the exam format and therefore supports the development that this student hopes for.

When asked to determine who should use the computer in class, the participants had various opinions. Almost a half opted out of taking stance on the statement “Every student has to have a computer at their use during English lessons”, with 25 % disagreeing and 20 % agreeing. Then again, only 8 % of the students thought that only the teacher should have a computer at her use in class, with 45 % of
the students disagreeing and 47% not being able to answer. Students therefore seem to suggest something in between these two options.

When moving on to the open question, phrased: “What, if anything, do you think should be done differently in language education once the digitalized Matriculation Examination steps in?”, two topics overwhelmingly dominate the answers: increasing the amount of ICT in teaching and adding emphasis on the development of oral skills. Of 93 participants, twelve answered that they did not know or opted out of answering and four students answered plainly “No” without further explanation. The remaining 78 students’ answers, however, almost exclusively deal with these two topics.

By far the most recurrent suggestion for improving language education was bringing more ICT into English education. More than half of the students who had suggestions for improving upper secondary English education, 42 out of 78 students, mentioned bringing more ICT into English teaching as the major improvement. Out of these, five students went as far as to propose that teaching and learning should be digitalized altogether, whereas the remaining 37 hoped for various levels of increase in the use of digital tools in teaching. In addition to these 42 students, there were three students that were excluded from the group who also mentioned increasing the use of computers in upper secondary school education, but did not relate this usage specifically to English lessons. Instead, they saw it as part of the school’s obligations or suggested an increase in the teaching of compulsory IT. Although the students were unclear about who should use the computer in class, they made clear their demand for more ICT-enriched English education.

Q71 “English teaching could also be digitalized almost entirely.”

Q72 “More digital material and teaching methods should be used.”

Q73 “It [English teaching] should become more digitalized and include more information analysis.”

Some students encouraged digitalizing teaching in more moderate amounts:

Q74 “Once in every period we’d have an IT lesson where we’d read online news and answer questions based on them.”

Q75 “A comment to question “In upper secondary teaching, it’s important to develop language skills in an online environment.” I disagree, at least it shouldn’t be done too much, because almost all young people spend time online, so they get a lot of practice for this in their free time as well. The important thing would be to concentrate on those things you can’t learn otherwise.”
Q75 expresses a concern that developing online skills in language classes will overshadow the learning of the language itself. It has been shown by Lakkala and Ilomäki (2013), however, that “those things you can’t learn otherwise” also include crucial analytic and media literacy skills which are not learned by simply spending time online. For the versatile development of these important skills, students need subject-specific guidance from the teacher. In the case of English, analyzing online content combines, in an ideal way, the development of language skills along with modern-day computer and media literacy.

Simultaneously, Q75 underlines that the activities in the language classroom should never be ones which the students could just as well complete at home. Language learning activity in class should, otherwise said, always differ considerably from independent work or free time activity completed at home. In a world where distant learning is becoming an increasingly valid possibility for learning of all kind, it is important that language lesson activities make use of the physical presence of students in the same classroom, the atmosphere created by this presence and the opportunity for developing both written and oral face-to-face communication skills. In this way, Q75 emphasizes the importance of suitable methodology in utilizing ICT-based teaching and learning tools in class. Computers are ideal in creating a platform for interaction independent of time and space, so that the interaction of English lessons can be stretched outside the classroom as well. This quality made possible by the internet makes ICT tools ideal for making homework activities more interactive, motivating and beneficial. In class, however, the teacher needs to pay special attention to finding ways which incorporate technology into language learning while simultaneously acknowledging and making use of class dynamics.

Another interesting proposal, also requested by Suomen Lukiolaisten Liitto (2013) in their list of improving ICT usage in Finnish upper secondary schools, was brought out by a student in the following answer:

Q76 “Making use of students’ and teachers’ experience on computer-based work.”

This student hoped for students to be more involved in the planning and completion of computer-based work, and as Suomen Lukiolaisten Liitto (2013) states, this kind of approach would better commit students to responsible and appropriate use of devices in the classroom. It would doubtless also bear favorable consequences to student motivation, diminish teacher-centeredness and encourage democracy in the classroom.
The second recurring improvement proposal was placing more emphasis on oral skills in English education. Students simply hoped for more speaking in the English lessons and more opportunities for recording their own speech.

Q77 “More effort should be put into the oral part than nowadays, so that people would know English in practice”
Q78 “more oral practice/communication with native English speakers”
Q79 “The oral part must be practiced. For example we’d give oral presentations during English lessons etc.”

As brought out in the discussion on the oral skills exercise in 4.1.4, the students view oral skills as increasingly important in English language use and measurement, which is logically observable in their views on English teaching as well.

4.3 Summary of results

To bring the discussion together, this section will summarize the results to the two research questions and make links to existing research.

The first research question considered the students’ perceptions of the digital Matriculation Examination for English, based on their experience and the example exercises. Answering this question included the discussion of the three sub-questions concerning the attitudes (section 4.1.1), experiences (sections 4.1.2-4.1.5) and suggestions (section 4.1.6) that the participants brought out in their answers. Out of these sub-questions, the second one gained the most weight, as the design of the research was seen to provide especially appropriate conditions for examining this question.

First of all, the participants were divided in their attitudes towards the importance of digitalizing the Matriculation Examination. A major part of them felt that this reform was a significant change, because they regarded the example exercises as distinguishably different from the paper versions they had seen. Adding the oral part to the test was considered important by the majority, regardless of the student’s attitude towards the digitalization. For their own Matriculation Examination, most of the students preferred the paper version, with 22% favoring the digital test format. This is less than reported by Hurme, Nummenmaa and Lehtinen (2013), who found 40% of their respondents willing to complete the foreign language exam using digital tools only.

Secondly, the students experienced their own test performance as slightly worse in a digital test surrounding than a paper test one. When asked about their
abilities to concentrate and bring out their language skills, they regarded these abilities as worse in the digital test. This may be due to the data gathering conditions as well as unaccustomedness to completing digital tests. However, these findings stress the importance of making the students comfortable with the completion of computer-based tests.

To consider some background factors affecting the experiences, female students were more opposed to the digitalization than men, expressed more reserved attitudes towards its use and evaluated their performance as poorer than male students. This is in line with the gender bias found in Suomen Lukiolaisten Liitto (2013). Contrary to the findings of Hurme, Nummenmaa and Lehtinen (2013), however, the students’ IT skills were not seen to affect the attitudes or experiences in this study. Out of all the answers concerning the students’ attitudes, experiences or abilities, only two correlated with the students’ estimates of IT skills: their self-assessed ability to work carefully and to make efficient use of the test time. Both students and teachers expressed their concern about the IT skills required in the Matriculation Examination in Lakkala and Ilomäki (2013). While this concern is somewhat mitigated by the results of this study, more research needs to be directed towards the issue.

To systematically review the good and bad sides of the Matriculation Examination as brought out by the students, we will return briefly to Bachman and Palmer’s (1996) test usefulness qualities, discussed in 2.2.5.

The participants in this study especially laid emphasis on the gains concerning the *authenticity* and *interactivity* of the test. The multiple-choice reading comprehension task was considered topical and described an everyday situation for the students. The format and layout of the reading comprehension texts was closer to what students are used to reading, which gave the impression of more authentic material. The oral exercise was specifically stressed, since it was seen to be crucial in working life, further studies and authentic language use situations in general. The students also acknowledged the new type of skills required in the test. The introduction of larger source material, a whole new oral task type and a digital test platform inspired multiple comments regarding the use of the new skills at use. The students thus saw the digital exam as more interactive and engaging than the paper version. Some students experienced trouble with e.g. working with a larger source material, which supports earlier findings that upper secondary school students
are lacking behind in the development of important modern-day media skills (Hautamäki et al., 2012; Kiili, 2012; Lakkala and Ilomäki, 2013; Pajarre, 2012).

Practicality received both enhancements and drawbacks, and also the impact of the digital examination was considered in positive as well as negative terms. Typing answers on the computer was felt to be practical and user-friendly, and especially choosing the correct alternative in multiple-choice tasks and drop-down menus was felt to be faster and handier than in the paper version. The oral test on the whole raised many practicality issues, but the participants mostly discussed them showing clear motivation to develop this task type. Other impractical aspects mentioned by students were having too many windows and programs open and the effect of technical problems. Also imagined future technical problems were considered, along with the possible negative impact of digital tests on teaching, learning and basic skills like hand-writing. The impact of the reformed test was also considered in positive terms: it was considered crucial to bring more ICT and oral skills teaching to English education.

The more technical qualities of reliability and construct validity were unsurprisingly not explicitly brought out by the students. Only the construct validity of the oral test type was questioned by a few students, who brought out their concern about oral test anxiety affecting the test performance. The students felt that bringing out their language skills and concentrating was comparable to or slightly more difficult than in paper-based tests, which implies that the students feel that they would, at the moment, perform better in paper-based tests.

As medication for the negative aspects described in the above discussion on test qualities, the participants hoped for a clear, user-friendly test system minimizing the effect of students’ unintentional technical errors (see Fulcher, 2003). Additionally, they hoped to see more versatile and real-life based task types in the digital Matriculation Examination, including e.g. videos, interviews and interactive and communicative oral tasks. While a minor part of the students advised against digitalizing the Matriculation Examination, others were convinced that with a little more work in developing both the test and teaching methods, the reform will be successful.

When moving on to the second research question, the relationship between the new digital exam and current English education was considered slightly unbalanced. While their English teaching was generally seen to develop the language
skills relevant for future studies and working life, the participants expressed divergent opinions on the ways in which ICT is utilized in current teaching. The differences between teaching groups were found to be great, which supports claims of conflicting teaching practices between schools and teachers (e.g. Anttonen, 2011; Lakkola and Ilomäki, 2013 and Suomen Lukiolaisten Liitto, 2013). The students were practically unanimous about the fact that the Matriculation Examination should affect the goals and contents of upper secondary English education.

On the device front, students remain hesitant: neither the 1:1 computing model nor the old teacher-only computing is directly accepted as the new standard. The participants of this study are then not ready for what Bebell and O’Dwyer (2010) predict to be the future norm in American classrooms. At the moment, the ideal for the students would perhaps be something in between. Rather than rooting for a technological revolution in their English teaching, the students seem to view ICT as a tool used to enrich and update the way of learning English at upper secondary school.

In the suggestions for improving upper secondary English teaching, two suggestions overpower. First of all the students forcefully argue for an increase in the use of ICT. These results echo the demands made by the students in Suomen Lukiolaisten Liitto (2013). The comments range from suggesting one lesson per period to take place in a computer laboratory to replacing textbooks with computers altogether. The second amelioration demanded by the students is more practice for oral skills. They markedly stress the importance of the ability to communicate orally in English, giving it equal or even higher weight than written skills. As methods for oral skills practice, students suggest presentations, real-life communication with English speakers and recording the students’ own speech.
5 Conclusion

This study examined the attitudes and perceptions of upper secondary school students on the digital Matriculation Examination for English as a foreign language. The participants completed example exercises of the upcoming exam format and then answered a questionnaire concerning this experience. Since the actual test software to be used in the digital Matriculation Examination was not available for use at the time of the data gathering, a learning environment called ViLLE was used to present the exercises to the students. In order to provide as full an understanding as possible of the students’ experiences and perceptions, a lot of attention and space was devoted to quotations of the students’ answers to the open questions of the questionnaire.

The students revealed curious attitudes towards the example exercises and expressed their appreciation of the authentic and interactive elements newly available in the test format. While hesitating on a range of practical notes and feeling unequipped to deal with some aspects of the new test experience, the students showed motivation to develop the digital Matriculation Examination and proved to be well worth listening to in the reform process. The participants evidently see this as a time of change in the upper secondary English education. Whether the pressure for change derives from the Matriculation Examination reform or the surrounding society, students visualize future upper secondary English education as the versatile learning of useful and relevant language skills in an authentic and interactive environment. Modern-day technology provides appropriate tools for this, and their use can be mastered and revised through combining the expertise of students and teachers alike. These tools can also be used to enhance oral practice, which the students feel should be emphasized a great deal more.

When assessing the validity of this study, it needs to be considered that the example exercises utilized in this study were not a precise reflection of future test tasks. Every attempt was made to make them as close to the plans so far published as possible. They were created by the researcher based on the examples published by the Matriculation Examination Board and received the approval of the Digabi project. Practical considerations also affected the realization of the data gathering, as only devices already available in the schools could be used. The exercises were completed as part of normal schoolwork and did not affect the grades of the students.
and, in some groups, it was difficult to create a test-like atmosphere guaranteeing everyone the chance to concentrate. The situation was therefore not ideal for evaluating their own test performance. Other issues affecting the results of the study are differences between teaching groups in the data gathering session, which was conducted separately for each group and could not be accurately replicated. The students therefore received slightly different instructions or support and were subject to different distraction factors. Overall, these results represent the students’ first-hand impressions of testing procedure which most of them were not used to.

Once the actual Matriculation Examination software is available for use, it would be important to repeat this kind of study with elements closer to the actual exam situation, with more time and more exercises. If the students had the chance to work with familiar test software, it would perhaps guide their attention to more specific elements of the test contents and exam situation. As this study only gave the students a brief look into four different task types, it would be beneficial to do more research on the particular task types. As demonstrated by the variety of perceptive suggestions in the data collected for this study, students are definitely worth listening to when developing testing systems and procedures. The rapidly changing educational environment of Finland provides ideal circumstances for compensating for the lack of context-based studies on the implementation of CALL and CBLT theories, acknowledged in e.g. Bax (2011), Huh and Hu (2005) and O’Dowd (2007).

This study was seemingly the first to attempt a realization of the oral Matriculation Examination task, and the results show that a lot more research remains to be done in developing this task type before its introduction in 2019. The efforts will, however, be worthwhile, as the clear majority of the students acknowledge the weight of oral skills in today’s language proficiency. Interesting viewpoints for further studies concerning the oral part would be e.g. test anxiety or computer anxiety in oral test situations (see e.g. Fengyang, 2012; Hewitt and Stephenson, 2012), or case studies testing different practical realization models.

The Matriculation Examination continues to hold an important position in future teaching and society, and preparing the students for its requirements therefore remains a central goal of upper secondary English teaching. Ideally, this goal would be reached surreptitiously through utilizing similar task types and pursuing shared goals, laid out clearly in the new National Curriculum. Along with enhancing the authenticity and interactiveness of Finland’s only form of standardized testing, this
reform is a chance for the examination to become permanently more open and transparent about its sources, processes and standards. Hopefully, a new kind of cooperation between students, teachers and test designers is also being established in the reform process to ensure that the exam preserves its relevance and high-quality measurement while constantly reflecting on the practicality and relevance of its contents and format.
References


<http://papert.org/articles/SituatingConstructionism.html> [9 June 2014]


Appendices

Appendix 1: Walker and White (2013:10): From CALL to TELL
Appendix 2: Summary of usefulness analysis for the computer-based Matriculation Examination
Appendix 3: Instructions given to students in the data gathering
Appendix 4: Example exercise scripts and screenshots
Appendix 5: Script of the instruction audio clip for the oral skills exercise
Appendix 6: Questionnaire
Appendix 7: Background information about participants
Appendix 8: Additional figures concerning the learning environment
Appendix 9: Enlargements of Pictures 2, 3, 4 and 5
Appendix 10: Translations of the students’ comments Q1-Q79
### Appendix 1: Walker and White (2013:10): From CALL to TELL

<table>
<thead>
<tr>
<th>Approach</th>
<th>Structural/Restricted CALL</th>
<th>Communicative/Open CALL</th>
<th>Integrative CALL</th>
<th>TELL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td>From mainframe to mobile</td>
<td>PCs</td>
<td>Multimedia, internet</td>
<td>Mobile devices, tablets, multiplayer games, virtual worlds</td>
</tr>
<tr>
<td><strong>English-teaching paradigm</strong></td>
<td>Grammar-translation and audio-lingual</td>
<td>Communicative language teaching</td>
<td>Content-based ESP/EAP</td>
<td>Communication, interaction</td>
</tr>
<tr>
<td><strong>View of language</strong></td>
<td>Structural (a formal structural system)</td>
<td>Cognitive (a mentally constructed system)</td>
<td>Socio-cognitive (developed in social interaction)</td>
<td>Structural, cognitive, socio-cognitive, adaptable</td>
</tr>
<tr>
<td><strong>Principal use of technology</strong></td>
<td>Drill and practice</td>
<td>Communicative exercises</td>
<td>Authentic discourse</td>
<td>Normalized</td>
</tr>
<tr>
<td><strong>Principal objective</strong></td>
<td>Accuracy</td>
<td>Fluency</td>
<td>Agency</td>
<td>Autonomy within community</td>
</tr>
<tr>
<td><strong>View on learning</strong></td>
<td>Behaviourism</td>
<td>Constructivism</td>
<td>Social constructivism/situated learning</td>
<td>Connectivism</td>
</tr>
<tr>
<td><strong>Role of technology</strong></td>
<td>Tutor</td>
<td>Tutee</td>
<td>Mediational tool</td>
<td>Environment, resource</td>
</tr>
</tbody>
</table>
### Appendix 2: Summary of usefulness analysis for the computer-based Matriculation Examination

<table>
<thead>
<tr>
<th>Quality</th>
<th>Assets</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reliability</strong></td>
<td>– Anonymity of rating procedure</td>
<td>– Comparability of PBLT &amp; CBLT Matriculation Examinations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– Creation of new rating procedure for oral skills tasks</td>
</tr>
<tr>
<td><strong>Construct validity</strong></td>
<td>– More versatile test tasks to measure overall linguistic capability</td>
<td>– Different degrees of computational competence may affect the student’s overall performance at least before the new curricula have been established</td>
</tr>
<tr>
<td><strong>Authenticity</strong></td>
<td>– Texts in a more authentic form</td>
<td>– Lost “bodily interaction” between thinking and writing by hand</td>
</tr>
<tr>
<td></td>
<td>– Audiovisual material made possible</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– More authentic exam mode</td>
<td></td>
</tr>
<tr>
<td><strong>Interactiveness</strong></td>
<td>– More modern learner strategies at use in the exam situation</td>
<td>– Lost “bodily interaction” between thinking and writing by hand</td>
</tr>
<tr>
<td></td>
<td>– Possibility to provide the chance to search online or communicate with others in the future</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– More source material for learners to operate</td>
<td></td>
</tr>
<tr>
<td><strong>Impact</strong></td>
<td>– Digitalization and modernization of teaching and learning (positive side)</td>
<td>– Worry about too much time spent with technology</td>
</tr>
<tr>
<td></td>
<td>– Increasing media and electronic literacy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– Matriculation becoming proof of computer literacy and oral skills competence</td>
<td></td>
</tr>
<tr>
<td><strong>Practicality</strong></td>
<td>– Instantaneous correction of multiple-choice items and short questions</td>
<td>– Expensive investments in schools</td>
</tr>
<tr>
<td></td>
<td>– Possible further technological advances in rating procedure</td>
<td>– risks concerning power outages or malfunctioning devices</td>
</tr>
<tr>
<td></td>
<td>– User-friendliness in writing long essays</td>
<td>– lack of IT expertise especially in small schools</td>
</tr>
<tr>
<td></td>
<td>– one-day exam</td>
<td>– need for constant updating of devices</td>
</tr>
</tbody>
</table>
Appendix 3: Instructions given to students in the data gathering

ViLLE-oppimisympäristön käyttöohjeet

1) Avaa Internet-selain (mieluiten Mozilla Firefox tai Google Chrome) ja mene osoitteeseen https://ville.utu.fi
2) Syötä seuraavat tunnukset kirjautumiskenttiin:

Sähköposti: [individual username for each student]
Salasana: [individual password for each student]

3) Kun olet kirjautunut sisään, valitse "Esimerkki-tehtäviä englannin sähköisestä ylioppilaskokeesta" -> "Example exercises"

4) Voit tehdä kolme ensimmäistä tehtävää haluamassasi järjestyksessä:
   1. Reading comprehension (multiple choice)
   2. Reading comprehension (open questions)
      a. Lue teksti ja vastaa suomeksi kysymyksiin niille esitetyihin kohtiin. Lähätä lopuksi vastaukset painamalla .
   3. Grammar and vocabulary
      a. Lue teksti ja valitse pudotusvalikoista tekstiin parhaiten sopivat sanat. Lähätä lopuksi vastaukset painamalla .

5) Viimeinen tehtävä "Oral exercise" tehdään kaikki yhdessä ohjeiden mukaan. Sen teko aloitetaan yhdessä sovittuun kellonaikaan ja sitä varten tarvitset kuulokkeet ja mikrofonin. Älä lähätä vastauksiasi arvioitavaksi ennen sitä!

Yleisiä ohjeita:


2. Lue tehtävien ohjeet kunnolla. Mikäli ohjeet eivät näy kokonaan, paina nuolinapista koko ohjekenttä näkyviin.


4. Lämmin kiitos tutkimukseen osallistumisesta!
Ohjeet suulliseen tehtävään (käydään läpi yhdessä tehtävään siirryttäessä)


2. Valitse ViLLEstä tehtävä 4. Oral exercise


4. Seuraa sitten kuulemiasi ohjeita ja reagoi suullisesti niiden mukaisesti.

5. Kun ohjetiedosto loppuu, paina äänitysnauhurista "Lopeta äänitys" ja tallenna luomasi tiedosto haluamallasi nimellä paikkaan, josta löydät sen (esim. Työpöydälle).

Appendix 4: Example exercise scripts and screenshots

Example exercise 1:

News site available at <http://newsylt.blogspot.fi/> [4 October 2014]
Screenshot of the current layout of ViLLE (altered since the data gathering)

Questions:

1. What is true of the news site?
   - Headlines deal with a variety of awards
   - The source of the texts is not made transparent
   - The site is obsessed with celebrity trivia
   - The awardees are of a questionable nature

2. According to the BBC article, what are the Academy Awards all about?
   - Creating a certain image of Hollywood
   - Ridiculing the human condition
   - Receiving funds for the movie industry
   - Lasting and unforgettable movie experiences

3. The BBC article claims that the movie 12 Years A Slave is
   - a tribute to Africa
   - a movie intended only for a specific population
   - a touching representation of global racial issues
   - a reminder of an era of exploitation

4. What is the focus of the Nobel Peace Prize article?
   - Comparing Putin and Obama
   - How Putin became eligible for nomination
   - Whether or not Putin will win the prize
   - The biases of Nobel laureates
5. Claire Vaye Watkins’ prize-winning book *Battleborn*

is a young author’s blunder
reflects the real world in an authentic way
is tracing the style of Dylan Thomas
exhibits the distinguished talent of the author

6. What does the news site include?

- Statistics about award winners
- Analysis on teen-age awardees
- Tables about viewer records
- Figures on Nobel nominees

*Example exercise 2:*

Source text:

**Dorothy Gilman, ‘Mrs. Pollifax’ Novelist, Dies at 88**

Dorothy Gilman, an espionage writer whose best-known heroine, Mrs. Pollifax, is very likely the only spy in literature to belong simultaneously to the Central Intelligence Agency and the local garden club, died on Thursday at her home in Rye Brook, N.Y. She was 88.

The cause was complications of Alzheimer’s disease, her family said.

In “The Unexpected Mrs. Pollifax” (1966), the first novel in what would be a 14-book series, Mrs. Gilman introduces Emily Pollifax, a 60-ish New Jersey widow bored by the compulsory round of tea and good works. In search of adventure, she offers her services to the C.I.A. — who, after all, is going to peg a suburban grandmother as a cold war secret agent? — and adventure she finds. In the course of the series, which concluded in 2000 with “Mrs. Pollifax Unveiled,” she fetches up in Mexico, Turkey, Thailand, China, Morocco, Sicily and elsewhere.

Clever, lucky and naïvely intrepid, Mrs. Pollifax employs common sense and a little karate to rescue the kidnapped; aid the resistance (when you are a suburban lady spy, a fashionable hat is ideal for concealing forged passports); and engage in all manner of cheery deception (when doing business with a malefactor who is expecting a can of plutonium, a can of peaches makes an excellent if short-term substitute).

Reviewers sometimes quibbled about the improbability of the novels’ basic premise. But the books proved popular with readers: in a genre in which women had long been young and sultry, Mrs. Pollifax, with her peril and petunias, made an irresistible, early feminist heroine.

The series was the basis of two movies, the feature film “*Mrs. Pollifax — Spy*” (1971), starring Rosalind Russell, and the telefilm “The Unexpected Mrs. Pollifax” (1999), starring Angela Lansbury.

The Mystery Writers of America named Mrs. Gilman its 2010 Grand Master.

Dorothy Edith Gilman was born in New Brunswick, N.J., on June 25, 1923; she decided on a writing career when she was still a child. Planning to write and illustrate books for children, she studied at the Pennsylvania Academy of the Fine Arts. Under her married name, Dorothy Gilman Butters, she began publishing children’s books in the late 1940s.
Mrs. Gilman’s marriage to Edgar A. Butters Jr. ended in divorce. She is survived by two sons, Christopher Butters and Jonathan Butters; and two grandchildren.

By the seventh Mrs. Pollifax novel, “Mrs. Pollifax and the Hong Kong Buddha,” published in 1985, Mrs. Gilman’s heroine has remarried. But for the most part, she is quite content to leave her husband at home for the duration of the series as she gads about the world, a paladin packing peaches.


Questions:

a) Kuka on Dorothy Gilman ja mitä hänen on tapahtunut?
b) Mistä syistä Emily Pollifax hakeutui C.I.A:n palvelukseen?
c) Mitä ominaisuuksiaan rouva Pollifax hyödyntää työssään?
d) Mitä tekstissä kerrotaan rouva Pollifaxin avioliitosta?

Example exercise 3:

Text:

The harbor, at first sight, I did not understand it. There were so many ships at anchor that their masts looked to be tangled impossibly; hundreds of them packed together so densely that it was as to as/that give the appearance of a vast, limbless forest rolling on the tides. Charlie and I threaded our way up the shoreline, and all around us was/were/there were/there is chaos. Men of every race and age rushing, shouting, pushing, fighting; cows and sheep had directed/direct/were directed/had been directed this way and that; horse-led wagons carried lumber and bricks up the mud-slick hill, and the sound of hammering and building echoed from the city for the sea/out to sea/into the sea/through the sea. There was laughter in the air, though/supposing/ despite/therefore it did not give me the impression of gaiety, but something more maniacal and evil wishing. Tub was nervous, and so was I. I had not seen anything closely/slightly/arguably/remotely like it and I wondered how we might possibly find one man in these labyrinthine streets and alleyways, where all was queer and dark and hidden.

"Help/Allow/Let/Show us search out Morris," I said. "He has already waited/being already waited/already waited/has already been waited weeks for us," said Charlie. "Another hour won't change anything." Of course my brother liked the atmosphere, and was not the least bit/a bit/the smallest/little uneasy.

*The Sisters Brothers*, Patrick DeWitt

Illustration of drop-down menu:
Appendix 5: **Script of the instruction audio clip for the oral skills exercise**

Created on the basis of the oral exercise example 1:

Website shown to students available at [http://digabioralskills.blogspot.fi/](http://digabioralskills.blogspot.fi/) [20 October 2014]

*Part A - reading*

Please read the sentences as you are instructed.

Please read sentence 1. [see below]

Now read sentence 8.

*Part B - repeat*

Please repeat each sentence that you hear. For example, if I say: “He had to go grocery shopping that day”, you say: "He had to go grocery shopping that day."

1) I was really embarrassed to hear what had happened.
2) It was a cloudy day with only a few glimpses of sunlight.
3) As he was illiterate, he had never learned to read or write.

*Part C - questions*

Now please just give a simple answer to the questions. For example, if I say "What is more like stone, cement or rubber?", you say "cement".

What month comes before October?

Is a banana a fruit or a vegetable?

Would you rather sit on a table or on a chair?

*Part D - story retelling*

You will hear a brief story. The story will only be spoken once, followed by a beep. When you hear the beep, you will have 30 seconds to retell the story in English. Try to retell as much of the story as you can, including the situation, characters, actions and ending. You will hear another beep at the end of the 30 seconds.

“Bill wanted to watch TV when he got home. His father said that first he had to walk the dog, then he had to clean his room. After he finished his chores, he could watch TV.”

*bEEP*

30 seconds

*bEEP*

*Part E - open questions*

You will hear a question about family life or personal choices. The question will be spoken twice, followed by a beep. When you hear the beep, you will have 40 seconds to answer the question. You will hear another beep at the end of the 40 seconds.

“Would you rather live in a large city or a small village?”

“Would you rather live in a large city or a small village?”

*bEEP*

40 seconds

*bEEP*

Thank you. Now stop recording and upload the file into the system.”

Sentences (listed below the video):

1. There was a lot of traffic on the main road.
2. Is there any real need for such laws?
3. Even serious scholars may at times be guilty of this.
4. I want to know everything about it!
5. Without language, we could hardly have created the world we know.
6. He said he'd received some terrible news.
7. It was fascinating to hear about it.
8. I never realized running could be so much fun.
9. She told her the truth about everything.
10. The company received tons of complaints during the first few weeks.
Appendix 6: Questionnaire

Pitkän englannin sähköinen ylioppilaskoe

Tämän kyselyn tarkoitus on selvittää opiskelijoiden mielipiteitä englannin sähköisestä ylioppilaskokeesta ja sen suhteesta kielten opetukseen.

Kysely koostuu monivalintakysymyksistä ja avoimista kysymyksistä ja vastaamiseen kuluu maksimissaan 15 minuuttia. Vastaathan jokaiseen kysymykseen totuudenmukaisesti ja huolella oman näkemyksesi mukaisesti. Juuri sinun mielipiteesi on kiinnostava!

Kysely täytetään nimettömästi ja vastaukset käsitellään luottamuksellisesti.

Kiitos kyselyyn vastaamisesta!

T: Sanna Tarvainen
sanna.tarvainen@helsinki.fi

Sähköinen ylioppilaskoe

täysin eri mieltä/eri mieltä/en osaa sanoa/samaa mieltä/täysin samaa mieltä

Tekninen toteutus

Minun oli hankala hahmottaa koejärjestelmiä.

Koejärjestelmä tuntui luotettavalta.

Koejärjestelmä oli erilainen kuin ohjelmat ja sovellukset, joita yleensä käytän.

Koejärjestelmän käyttö oli helppo omaksua.

Olen käyttänyt samantyylistä järjestelmiä aikaisemmin.

Kokemuksesi sähköisen ylioppilaskokeen esimerkkitehtävistä

Englannin ylioppilaskokeen sähköistäminen on tärkeä uudistus.

Englannin sähköinen ylioppilaskoe vaikuttaa samantyyppiseltä kuin paperinen ylioppilaskoe.

Esimerkkitehtävät kuvastavat tilanteita, joita voisim kohdata koetilanteen ulkopuolella.
Englannin sähköinen ylioppilaskoe testaa kielitaitoani monipuolisesti.

Olen iloinen, että saan suorittaa ylioppilastutkinnon ilman suullista osuutta.

Englannin sähköinen ylioppilaskoe vaikutaa vaikeammalta kuin näkemäni paperiset ylioppilaskokeet.

Jos saisin valita, tekisin englannin ylioppilaskokeen sähköisessä muodossa.

On tärkeää testata suullista osaamista ylioppilaskokeessa.

**Kuvaile vapaasti 1-3 adjektiivilla (esim. Kilpikonna on mielestäni: vihreä, limainen, sympaattinen)**

**Kuvaus (1-3 sanaa)**

Luetunymmärtämistä mittaava monivalintatehtävä (uutissivusto) oli mielestäni:

Luetunymmärtämistä mittaava aukkotehtävä (rouva Gillman) oli mielestäni:

Sanasto- ja rakennetehtävä oli mielestäni:

Suullinen tehtävä oli mielestäni:

**Sähköisen ja paperisen kokeen vertailu**

*huomattavasti huonommin kuin paperilla/ huonommin kuin paperilla/ yhtä hyvin/huonosti kuin paperilla/ paremmin kuin paperilla/ huomattavasti paremmin kuin paperilla*

**Sähköisessä koetilanteessa**

pystyn tuomaan esille kielen osaamiseni

pystyn työskentelemään huolellisesti
pystyn keskittymään työhöni

pystyn käyttämään koesuoritukseen annetun ajan tehokkaasti

koen oppivani koetta tehdessäni

**Sähköinen ylioppilaskoe ja englannin opetus**

täysin eri mieltä/eri mieltä/en osaa sanoa/samaa mieltä/täysin samaa mieltä

**Lukion aikana saamani englannin opetus…**

valmistaa minua hyvin tämäntyyppiseen kielikokeeseen.

hyödyntää teknologiaa opetuksessa tehokkaasti.

kehittää tulevissa opinnoissa ja työelämässä tarvittavia kielitaitoja.

hyödyntää oppimateriaaleja, joissa on esimerkkitehtävien tyylisiä tehtäviä.

painottaa tasapainoisesti suullista ja kirjallista kielitaitoa.

sisältää oman puheen äänittämistä ja kuuntelemista.

kehittää kielenkäyttötilanteissa tarvittavia teknisiä taitoja.

hyödyntää teknologiaa opetuksessa opettajasta riippumatta.

**Lukion englanninopetuksessa on tärkeää**

harjoitella ylioppilaskokeen tehtävätyyppejä.

kehittää kielitaitoa Internet-ympäristössä.

imitoida tilanteita, joita voisin kohdata koulun ulkopuolisessa elämässä.
olla vuorovaikutuksessa muiden englanninpuhujien kanssa.

että opiskelija itse käyttää tietokonetta tunnilla.

Miten englannin opetusta lukiossa pitäisi mielestäsi kehittää?

Lukion englannin opetuksen pitää valmentaa opiskelijaa ylioppilaskirjoituksiin.

Englannin opetuksen tulee pysyä samanlaisena kuin saamani englannin opetus.

Ylioppilaskokeen vaatimusten kuuluu vaikuttaa lukion englannin opetuksen sisällöön.

Lukioenglannin oppikirjat on korvattava sähköisillä oppimateriaaleilla ylioppilastutkinnon sähköistyessä.

Jokaisella opiskelijalla on oltava käytössään tietokone englannin oppitunneilla.

Englannin oppitunneilla on harjoiteltava samoja tehtävätyypejä kuin ylioppilaskirjoituksissa.

Oppikirjoja on käytettävä englannin opetuksessa.

Englannin oppitunnilla vain opettajalla on oltava tietokone käytössään.

Avoimet kysymykset

Mikä esimerkkitehtävissä oli hyvää?
Mikä esimerkkitehtävissä oli huonoa?
Miten englannin sähköstä ylioppilaskoetta voisi mielestäsi kehittää?
Pitäisikö englanninopetusta mielestäsi jotenkin kehittää, kun ylioppilaskirjoitukset sähköistyvät?
Miten?
Tähän voit halutessasi vielä vapaasti kommentoida esimerkkitehtäviä, kyselytä tai sähköstä ylioppilaskoetta.

Taustatiedot

Sukupuoli
Esimerkkikehtävissä käyttämäsi ViILLE-käyttäjätunnus
Arvioi omaa englannin osaamistasi kouluarvosanalla 4-10
Arvioi kouluarvosanalla 4-10, kuinka mieluisa oppiaine englanti sinulle on
Arvioi omat tietotekniset taitot kouluarvosanalla 4-10
### Appendix 7: Background information about participants

#### Gender

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man</td>
<td>55</td>
<td>58.5</td>
<td>58.5</td>
<td>58.5</td>
</tr>
<tr>
<td>Woman</td>
<td>39</td>
<td>41.5</td>
<td>41.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

**Gender of participants**

#### Estimate of own English skills

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 to 6, poor</td>
<td>7</td>
<td>7.4</td>
<td>7.4</td>
<td>7.4</td>
</tr>
<tr>
<td>7, satisfactory</td>
<td>17</td>
<td>18.1</td>
<td>18.1</td>
<td>25.5</td>
</tr>
<tr>
<td>8, good</td>
<td>31</td>
<td>33.0</td>
<td>33.0</td>
<td>58.5</td>
</tr>
<tr>
<td>9, very good</td>
<td>31</td>
<td>33.0</td>
<td>33.0</td>
<td>91.5</td>
</tr>
<tr>
<td>10, excellent</td>
<td>8</td>
<td>8.5</td>
<td>8.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

**Participants’ estimates of own English skills**

#### Estimate of English agreeability

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 to 6, poor</td>
<td>11</td>
<td>11.7</td>
<td>11.7</td>
<td>11.7</td>
</tr>
<tr>
<td>7, satisfactory</td>
<td>8</td>
<td>8.5</td>
<td>8.5</td>
<td>20.2</td>
</tr>
<tr>
<td>8, good</td>
<td>36</td>
<td>38.3</td>
<td>38.3</td>
<td>58.5</td>
</tr>
<tr>
<td>9, very good</td>
<td>25</td>
<td>26.6</td>
<td>26.6</td>
<td>85.1</td>
</tr>
<tr>
<td>10, excellent</td>
<td>14</td>
<td>14.9</td>
<td>14.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

**Participants’ estimates of English agreeability**

#### Estimate of own IT skills

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 to 6, poor</td>
<td>10</td>
<td>10.6</td>
<td>10.6</td>
<td>10.6</td>
</tr>
<tr>
<td>7, satisfactory</td>
<td>18</td>
<td>19.1</td>
<td>19.1</td>
<td>29.8</td>
</tr>
<tr>
<td>8, good</td>
<td>28</td>
<td>29.8</td>
<td>29.8</td>
<td>59.6</td>
</tr>
<tr>
<td>9, very good</td>
<td>30</td>
<td>31.9</td>
<td>31.9</td>
<td>91.5</td>
</tr>
<tr>
<td>10, excellent</td>
<td>8</td>
<td>8.5</td>
<td>8.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

**Participants’ estimates of own IT skills**
Appendix 8: Additional figures concerning the learning environment

**Learning to use the test software was easy.**

**The test software was different from the programs and applications that I normally use.**
Appendix 9: Enlargements of Pictures 2, 3, 4 and 5
### Appendix 10: Translations of the students’ comments

| Q1 | “I think the Matriculation Examination for English should not be digitalized! There are too many risks related to technology, which may ruin the whole test performance. An oral part could still be added to the test.” | “Mielestäni englannin yliopilkaskokeita ei pitäisi sähköistää! On liikaa riskejä liittyen teknologiaan, jotka saattavat pilata koko koestuoriutoksen. Kokeeseen voisi silti lisätä suullisen osan.” |
| Q2 | “When I opened the exercises, I noticed that regardless of the instructions, I wasn’t sure what to do.” | “Avattessani tehtäviä, huomasin, että ohjeista huolimatta en ollut varma miten toimia.” |
| Q3 | “The graphics were weird somehow childish to some extent” | “Grafiikat olivat oikeutettavissa, mutta niistä helpommin ymmärtävän.” |
| Q4 | “The website didn’t seem official.” | “Sivusto ei vaikuttanut viralliseksi.” |
| Q5 | “Unclearness and uncertainty about whether the answers are going where they should go.” | “Epäselvyys ja epävarmuus siitä menevätkö vastaukset perille.” |
| Q6 | “Making the system simpler and clearer is important.” | “Systeemin yksinkertaistaminen ja selkeytäminen on tärkeää.” |
| Q7 | “the site could be made clearer also with regards to its looks, and it could be made to feel more secure” | “Sivustosta voisi tehdä selvimmin myös ulkonäkönsä puolesta, ja siitä voisi tehdä varmemman tuntuisen.” |
| Q8 | “There is still a need for development in the system and its graphical user interface (gui).” | “Järjestelmässä ja sen käyttöliittymässä (gui) on vielä kehittämistä.” |
| Q9 | “The text exercises could have a specific button you could press every time you’re unsure of some answer, so that when you’re checking, you would remember where you need to pay more attention.” | “Tekstitehtävissä voisi olla joku nappi mitä voisi painaa, aina kun on epävarmaa jostain vastauksesta, jotta tarkastelemma muistaa mihin täytyy kiinnittää enemmän huomiota.” |
| Q10 | “The “end test” button should be moved somewhere where you can’t click it accidentally so easily. -- You should be able to edit your answers until you have sent the whole exam and logged out of the service.” | “Lopeta koe nappi johonkin paikkaan missä siihen ei osu vahingossa niin helposti. -- Vastauksiapitäisi saada muokata niin kauan kunnes on lähettänyt koko valmiin kokeen ja kirjautunut ulos palvelusta.” |
| Q11 | “The “end test” button could be at the very bottom of the main menu and it could still ask for verification that the student wants to end the test. An x would be needed in the corner of multiple-choice questions, so that the windows could be closed without having to choose an answer.” | “Lopeta koe-nappi olisi ollut päälänsä ihan alhaalla, ja vielä varmistaa, että oppilas haluaisi lopettaa kokeen.Rasti tarvittaisiin monivalintakysymyksiin ikkunoiden nurkkaan, jotta ikkanat voisi sulkea vaihtoehtoa.” |
| Q12 | “My attention was also caught by the time that kept running in the left-hand corner of the screen. Doesn’t a specific time frame just cause stress to the student, when you can’t decide yourself what to do first etc?” | “Minua kiinnosti myös aika joka pyöri ruudun alhaalla. Eikö tietty aikamäärä vain stressaa oppilasta, kun ei voi itse päättää mitä tekee ensiksi jne?” |
| Q13 | “The software could be simpler and it could proceed like “on its own” so that the student doesn’t have to keep clicking new places, because it makes you confused” | “Ohjelma voisi olla selkeämpä ja mennä ns.” |
| Q14 | “Bigger and thicker fonts, for someone with dyslexia it’s extremely difficult to follow the text without color foils and without high lightings or underlinings.” | “Isommat ja paksummat fontit, lukihäiriöiselle erittäin vaikeaa seurata tekstiä ilman värikalvoja ja ilman yli/alleviivauksia.” |
| Q15 | “Some were similar to the tasks in the paper version but did not include flipping through papers etc.” | “Osaa muistutti paperiversion tehtäviä mutteivät sisältäneet paperien selailua yms.” |
| Q16 | “The structure of the exercises was nearly identical with that of the paper-based Matriculation Examination.” | “Tehtävät olivat rakenteeltaan laki yhtenevät paperisen YO-kokeen kanssa.” |
| Q17 | “They [the exercises] were similar to the normal exam.” | “Ne olivat samankaltaisia kuin normaaliassa kokeessa.” |
| Q18 | “They were difficult because there were so many links and tabs. It made me feel I know a lot less than I had thought.” | “Ne olivat hankalit, kun linkkejä ja välilehtiä oli monta. Sai tuntemaan että olen paljon vähemmän mitä olen huollut.” |
| Q19 | “Interesting texts, for which well-suited questions, joy of learning” | “Mielenkuntoisia teoksia, joihin mielekkää kysymyksiä, oppimisen iloa” |
| Q20 | “They were about topical issues and topics that interest people of our age.” | “Ne olivat ajankohtaisista aiheista ja meidän ikäisiä kiinnostavia aiheita.” |
| Q21 | “The texts were interesting, and not too dry expository texts. Their topicality, too, made them easier to understand.” | “Tekstit olivat mielenkiintoisia, eivätkä turhan kuvia asiakkeista. Niiden ajankohtaisuuskin teki niistä helpommin ymmärrettävän.” |
Q22 “They were more about the real world for example news were used in the reading comprehension exercises”  
"Koskivat ememmän ulkomaalmaa esimerkiksi luutunymmärtämisissä oli käytetty uutisia.”

Q23 “Reading comprehension exercises were good, because they measured your skills in reality.”  
"Luutun ymmäröimistehotiettävät olivat hyviä, koska ne mittasivat taitoa oikeasti.”

Q24 “In the online news you had to use your brain a little when they didn’t tell you directly where to find the answer. They were however quite limited in range, but it’s likely that the future Matriculation Exam will include more versatile questions.”  
"Nettiuutisissa piti vähän käyttää aivojansa kun ei suoraan kerrottu, että mistä vastaus löytyy. Ne olivat tosin melko suppeita, mutta todennäköisesti mahdollisessa yo-kokeessa on monipuolisuus kysymyksissä.”

Q25 “It was a little difficult to follow which question dealt with which article (in the first task).”  
"Oli hieman vaikea seurata mihin artikkeliin kukaan kysymys kuului (ensimmäisessä harjoitustehottävissä).”

Q26 “I think the first exercise had too many tabs, which may confuse the student.”  
"Mielestäni ensimmäisessä tehtävissä oli liikaa välilehtejä, mikä saattaa sekoittaa oppilasta.”

Q27 “The menu in the multiple-choice task for reading comprehension was weird and there were a bit too many sites to open”  
"Luutunymmärtämisenseen monivalintatehtävän valikko oli omituinen ja avattavia sivustoja oli vähän liikaa.”

Q28 “In the reading comprehension the pictures were disturbing concentration and they could be left out.”  
"Luutun ymmäröimistehottävissä kuvat häiritisivät keskitymistä ja ne voisi jättää pois.”

Q29 “Answering was easy and fast”  
"Helppo ja nopea vastata”

Q30 “You could write on the computer. If you made mistakes it was easy to just press delete.”  
"Pystyi kirjoittamaan koneella. Jos tuli virheitä ei tarvinnut muuta kuin painaa deleteä”

Q31 “In the reading comprehension the articles would be displayed in an easier way (like in exercise 2)”  
"Luutun ymmäröimisessä artikkelit olisivat helpommin esillä (niinkuin tehtävissä 2)”

Q32 “answering did not take time, it was easy to just click and it was nice when in the cloze exercise in the grammar section the completed text could be read in its entirety”  
"Vastaukseen ei kulunut aikaa, oli helppo vaan klikata ja oli mukavaa kun aukkotäydennystekstien grammar osiossa sai luettaa valmiin tekstin kokonaisuutena”

Q33 “it was easy to change the choices in the cloze exercise, clarity in the cloze exercise, when you did not have to check to see whether the ball is at the right number”  
"aukkotäytävän valintoja helppo vaihata, selkeys aukkotehtävissä, kun ei tarvinnut katsoa, onko pallo oikean numeron kohdalla,”

Q34 “the listening exercise, because it was really disturbing when others talked at the same time and you couldn’t hear the questions properly”  
"kuuntelu tehtävä, kun siinä häiristi tosi paljon kun muut puhu samaa aikaa eikä kysymyksiä kuullut kunnolla.”

Q35 “The oral exam should be arranged - in separate rooms or using noise-cancelling headphones.”  
"Suullinen koe tulisi järjestää joko erillisissä huoneissa tai vastamelukuulokkeilla.”

Q36 “It was painful listening to the instructions for the oral exercise from crappy earphones.”  
"Oli tuskalli kuulokkeista kuulokkeistä.”

Q37 “I do not think the oral part was good. Headphones should cut off noise notably better in order for it to work.”  
"Suullinen osio ei ollut mielestäni hyvä.  Kuulokkeiden pitäisi estää melua huomattavasti paremmin jotta se toimisi.”

Q38 “In addition, the program could have its own recording facility and noise-cancelling headphones for the oral part. (so that cheating wouldn’t be as easy)”  
" lisäksi ohjelmassa olisi hyvä olla oma äänitystehotietävät ja äänity lähipäätämättömät kuulokkeet suullista osaa varten. (jollei lunttaaminen olisi niin helppoa)”

Q39 “the oral exercise was somehow distressing”  
"suullinen tehtävä oli jotenkin ahdistava”

Q40 “The oral part was not realized in a very practical way. The noise in the classroom made me forget for example how I had started the sentence or what I was about to say and also made hearing difficult. In the actual Matriculation Examination it should be better realized. It should also be considered that the student might not succeed as well in the oral part if they have a shy temper and do not enjoy performing. When you are anxious in the test situation, blood hums in your ears and your mouth is dry, which makes talking more difficult.”  

Q41 “You can’t do something like this if the earphones are that bad, go and buy a f***** 10 E headset at Prisma and you’ll hear better than out of those cancer gimmicks”  
"Ei tällä voi tehdä jos on noin huonot kuulokkeet pil, ostakaa prismasta v f****** E headset ja siitä kuuluu paremmin ku noista syöpä-rohjoista.”

Q42 “The oral test was nice :)”  
"suullinen koe oli kiva :)

Q43 “The oral test was good in that you got to show your”  
"Puhekoe oli siitä hyvä, että siinä pääsi näyttämään
| Q44 | “I think testing oral language capacity was EXTREMELY GOOD! Indeed, that’s exactly why we learn languages so we could use it in real life as well.” | “suullisen kielitaidon testaaminen oli mielestäni ERITTÄIN HYVÄ! Siksihän me kielä juuri opimme että opimme myös käyttämään sitä oikeassa elämässä.” |
| Q45 | “unaccustomedness to completing an oral test also evoked confusion but had the task type been practiced beforehand, it would be less confusing” | “tottumattomuus suullisen kokeen tekemiseen aiheutti myös hämennystä mutta mikäli tehtävätyyppiä olisi harjoiteltu aikaisemmin olisi se vähemmän hämmentävä.” |
| Q46 | “In the oral part you can’t correct yourself if you get it wrong.” | “Suullisessa osiossa ei voi korjata jos menee väärin.” |
| Q47 | “-- in the oral part the result can be remarkably worsened by anxiety in the test situation, which may cause for example stuttering.” | “Toisaalta suullisessa osiossa tulosta voi huonontaa huononn avasti kokeessa oleva jännitys, joka saattaa aiheuttaa esim. äntyttämistä.” |
| Q48 | “The oral part is however almost the most important thing to know about the English language and you need it the most. Its testing should be increased and developed a lot.” | “Suullinen osuus on kuitenkin melkein tärkein osa englannin kielellä ja sitä tarvitsee eniten. Sen testaamista tulisi lisätä ja kehittää paljon.” |
| Q49 | “Organizing the oral exam was poor in the test situation. For sure, hearing others speak and the resulting cheating and disturbances will probably be minimized in the real exam situation, but in this test situation hearing others speaking was very disturbing to me. often I couldn’t hear the instructions, and it was difficult to repeat example sentences, when I’d only heard half of them.” | “suullisen kokeen organisointi oli testitilanteessa heikkoa. Toki todellisessa koetilanteessa muiden puheen kuuleminen ja siitä aiheutuvat lunttaamiset ja häiriöt varmaankin minimoidaan, mutta testitilanteessa koin muiden puheen kuulumisen hyvin häiritseväksi, ohjeet järven usein kuulumatta, ja esimerkkialaista olivat vaikeaa toistaa, kun oli itse kuullut niistä vain puolet.” |
| Q50 | “I think there should be more listening comprehension and discussion situations” | “Mielestäni pitäisi kulunymmärtämästä olla enemmän ja keskustelutilanteita” |
| Q51 | “I wonder if the oral part is sensible.” | “Onko suullinen osuus järkevä?” |
| Q52 | “The oral exam was impractical and needs improvement. I myself feel that the oral exercise is unnecessary.” | “Suullinen tehtävä epäkäytännöllinen ja parannusta vaatava. Itse koen suullisen tehtävän tarpeettomaksi.” |
| Q53 | “It was handy reading off the screen. Nicer to complete than the paper version.” | “Oli kätevää lukea ruudulta. Kivempi tehdä kuin paprista.” |
| Q54 | “rather handy and you don’t have to write by hand” | “aika näppärä eikä tarvi kirjoittaa käsin” |
| Q55 | “It’s faster to work on the computer and for example correcting mistakes is easier.” | “Tietokoneella ne on nopeampi tehdä ja esimerkkikisikin virheiden korjaaminen on helpompaa.” |
| Q56 | “It was easy to kind of ‘erase’ if you answered wrong.” | “Tehtävissä oli ns. helppo ”kumittaa”, jos vastasi väärin.” |
| Q57 | “It’s overall worse to do e exercises on the computer.” | “tietokoneella on ylipääätänsä huonompi tehdä tehtävää” |
| Q58 | “The realization was extremely old-fashioned but the idea was good” | “voteutus oli erittäin vanhanaikainen mutta idea hyvä” |
| Q59 | “I think the test could include something that corresponds to the earlier listening comprehension section (though shorter), because I guess at some lectures you need to be able to listen closely to long speeches as well.” | “mielestäni kokeessa voisi olla jokin entistä kuulunnymmärtämäksiksi vastaava osio (tosin lyhyempiän), sillä kai jossain luennoillan pitää osata kuunnella tarkasti pitkähän puheita.” |
| Q60 | “an oral communication exercise where the aim is to respond understandably to different situations which you could face on the street for example.” | “suullinen kommunikaatio tehtävä jossa tarkoituksena on vastata ymmärtettävästi erilaisiin tilanteisiin joihin voi törätä kaduilla esimerkkiksi.” |
| Q61 | “I think moving picture could also be included e.g. TED interviews and videos” | “Mielestäni mukaan voisi ottaa liikkuva kuvaa esim TED haastatteluja ja videoita” |
| Q62 | “Only the answers on the computer” | “Pelkät vastaukset tietokoneelle.” |
| Q63 | “The Matriculation Examination should not be digitalized. Computers can crash any time, and it does not feel as natural as writing with your own little hands. On top of everything, the contents of the examination will change for the worse along with the digitalization.” | “Ylioppilaskokeita ei pitäisi sähköistää. Koneet voivat kaatumia milloinkaan, eikä se tunnu niin luotettavaksi kuin omissa kädensä kirjoittaminen. Kaiken lisäksi kokeiden sisältö muuttuu sähköistymisen vuoksi.” |
| Q64 | “If all Matriculation Examinations are digitalized, people will lose the ability to write by hand.” | “Jos kaikki jo kirjoitukset sähköistyy, ihmisiä katoaa kyky kirjoittaa käsin.” |
| Q65 | “takes a little getting used to” | “vaatii hiukan totuttelua” |
| Q66 | “For now, the paper version seems more natural, even though it’s a question of getting used to.” | “Nyt paperiversio tuntuu mielestäni luontevammalta, vaikka se onkin
<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q67</td>
<td>“[it should be developed] towards a bit more familiar direction, but maintaining some features, such as the demand for heightened information analysis.” “Kehittää hieman tutumpaan suuntaan, mutta silti säilyttäen joitain piirteitä, kuten korostuneen tiedonerittelyvyn vaatimuksesta.”</td>
</tr>
<tr>
<td>Q68</td>
<td>“When students are familiarized with the digitalized Matriculation Examination and the exam tasks are developed a bit further and the system made better, the reform will surely be successful.” “Kun oppilaita tututetaan sähköisen ylppäreiden ja kokeen tehtävien välillä hieman kehitetään ja järjestelmä parannellaan, uudistus kyllä onnistuu.”</td>
</tr>
<tr>
<td>Q69</td>
<td>“The test was otherwise good, but it felt a bit weird compared to the paper version.” “Koe oli muuten hyvä, mutta tuntui vähän uudelta verrattuna paperiseen versioon.”</td>
</tr>
<tr>
<td>Q70</td>
<td>“We shouldn’t concentrate simply on studying the Matriculation Examination all the time but we should put more effort into learning to manage with your own English language if you really need to use it.” “Ei keskittyttäisi kokosojan vain ylppäreiden opiskelun vaan panostettaisiin enemmän että pärjätään omalla englanninkielellä jos sitä oikeasti joutuu käyttämään.”</td>
</tr>
<tr>
<td>Q71</td>
<td>“English teaching could also be digitalized almost entirely.” “englannin opetustulokset voitaisiin siirtää koneelle melkein kokonaan.”</td>
</tr>
<tr>
<td>Q72</td>
<td>“More digital material and teaching methods should be used.” “Pitäisi käyttää enemmän sähköisiä materiaaleja ja opetustapoja.”</td>
</tr>
<tr>
<td>Q73</td>
<td>“[English teaching] should become more digitalized and include more information analysis.” “Sen pitäisi muuttaa sähköisemmäksi ja tietoa erittelemään.”</td>
</tr>
<tr>
<td>Q74</td>
<td>“Once in every period we’d have an IT lesson where we’d read online news and answer questions based on them.” “Kerran jaksossa ATK-tunti, jossa luettaisiin nettiuutisia ja vastattaisiin kysymyksiin niiden pohjalta.”</td>
</tr>
<tr>
<td>Q75</td>
<td>“A comment to question “In upper secondary teaching, it’s important to develop language skills in an online environment.” I disagree, at least it shouldn’t be done too much, because almost all young people spend time online, so they get a lot of practice for this in their free time as well. The important thing would be to concentrate on those things you can’t learn otherwise.” “Kommentti kysymykseen ”Lukion opetuksessa on tärkeää kehittää kielitaitoa Internet-ympäristössä.” Olen eri mieltä, ainakaan sitä ei tulisi tehdä kovin paljon, sillä lähes kaikki nuoret viettävät aikaa internetissä, jolloin kyseiselle osa-alueelle saa harjoitusta paljon vapaa-ajanakin. Tärkeää olisi keskittyä niihin asioihin, joita ei muuten opisi.”</td>
</tr>
<tr>
<td>Q76</td>
<td>“Making use of students’ and teachers’ experience on computer-based work.” “Oppilaiden ja opettajien kokemusta sähköisestä työskentelyestä hyödyntämällä.”</td>
</tr>
<tr>
<td>Q77</td>
<td>“More effort should be put into the oral part than nowadays, so that people would know English in practice” “Suulliseen osuteen pitäisi panostaa enemmän kuin nykyään, että englantia osataan käytännössä.”</td>
</tr>
<tr>
<td>Q78</td>
<td>“more oral practice/communication with native English speakers” “enemmän suullista harjoittelua/kommunikointia englantia äidinkielisenä puhuvien kanssa”</td>
</tr>
<tr>
<td>Q79</td>
<td>“The oral part must be practiced. For example we’d give oral presentations during English lessons etc.” “Suullista osuutta täytyy harjoitella. Esim. Pidettäisiin englannin tunneilla suullisia esitelmiä tms.”</td>
</tr>
</tbody>
</table>