



**METROPOLIA UNIVERSITY OF APPLIED SCIENCES**

**FACULTY OF TECHNOLOGY**

**Bachelor's Degree in Industrial Management**

**Bachelor's Thesis**

**DIAGNOSTIC LANGUAGE TEST RESULTS AND NEED FOR PREPARATORY  
LANGUAGE STUDY**

**Author: Miika Kuparinen  
Instructor: Marjatta Huhta**

**Approved: 6. 6. 2011**

**Marjatta Huhta, DSc (Tech)**



## **PREFACE**

This thesis was an interesting journey, that faced several challenges and the deadlines came and went like the seasons and tides during the process. Now that it is finished I have a sense of pride but also a feeling of accomplishment after finally reaching the goal I aimed for so long ago entering Metropolia as a new student.

But I could not have done this just by myself, and wish to use this opportunity to thank the people making this study possible and offering me an access to the data regarding their Universities of Applied Sciences, Ritva Ala-Louko of Rovaniemi AMK, Irene Hyrkstedt from Savonia, Joni Sallila from SAMK and Linda Saukko-Rauta from JAMK, I would like to thank you personally for making this crucial data available for me. Also would like to express my gratitude towards the other teachers who kindly took the time and answered my e-mail query regarding the use of diagnostic testing in different UAS's across Finland. Finally I would like to thank my instructor Marjatta Huhta for all the support, guidance and general help, I needed those and without them this thesis would not have been finished.

Helsinki, June 6, 2011

Miika Kuparinen

## ABSTRACT

<b>Name:</b> Miika Kuparinen	
<b>Title:</b> Diagnostic Language Test Results and Need for Preparatory Language Study	
<b>Date:</b> April 2, 2009	<b>Number of pages:</b> 38
<b>Degree Programme:</b> Industrial Management	
<b>Instructor:</b> Marjatta Huhta, DSc (Tech)	
<p>The intention of this Bachelors thesis study was to answer two questions that had arose concerning the diagnostic language test held in Metropolia University of Applied Sciences; do students of Metropolia possess the skill levels required in foreign languages, and whether or not there is a need for a diagnostic test at all.</p> <p>Theoretical background of this study discusses the legislative demands for language skills for students graduating from UAS, and the practical applications of these levels, and also do a quick overview on the diagnostic testing in general, to see if and what testing is used in similar situations.</p> <p>Practical part of the study goes through the results from four years, a total data sample of over 6500 students in Swedish and 5500 in English, and tries to draw conclusions on the questions asked, and look for positive correlation between grades received in previous educational levels and the result gotten in the diagnostic test to determine the need for said testing.</p> <p>As majority of the students pass the test in both languages, the results gotten from this study also support the view that majority of students cope with the test as their previous grades would suggest, and testing of this group is not meaningful. But there is also a group of students of which a definitive conclusion cannot be made as their results vary between passing and failing, and a big question is if the diagnostic testing is deemed unnecessary which category would this group fall, passing or failing.</p>	
<b>Key words:</b> language study, diagnostic testing	

## OPINNÄYTETYÖN TIIVISTELMÄ

<b>Työn tekijä:</b> Miika Kuparinen	
<b>Työn nimi:</b> Diagnostic Language Test Results and Need for Preparatory Language Study	
<b>Päivämäärä:</b> 6.6.2011	<b>Sivumäärä:</b> 38 s
<b>Koulutusohjelma:</b> Tuotantotalous	<b>Ammatillinen suuntautuminen:</b> Tuotantotalous
<b>Työn ohjaaja:</b> TkT Marjatta Huhta	
<p>Tässä Insinööriyössä keskityttiin ratkaisemaan kaksi esillenoussutta kysymystä koskien Metropolian organisoimaa vieraiden kielten diagnostista testiä; omaavatko Metropoliasa opiskelut aloittavat opiskelijat vaadittavan kielitaidon, ja onko kyseisenkaltaiselle diagnostiselle testaukselle todellista tarvetta?</p> <p>Työssä käydään läpi lainsäädännöllisiä vaatimuksia joita on asetettu ammattikorkeakouluopiskelijalta vaadittavasta kielitaidon tasosta, sekä vertaillaan muita käytettyjä kielitaidon testausmenetelmiä Metropolian oman testin lisäksi.</p> <p>Työtä varten on kerätty tiedot neljältä vuodelta Metropoliasa aloittaneista opiskelijoista kahdesta kielestä, ruotsista ja englannista. Ruotsin otos on yli 6500 opiskelijaa ja englannin yli 5500 neljältä vuodelta. Tuloksista pyritään saamaan hyvä yleiskuva opiskelijoiden vallitsevasta kielitaidosta ja sen riittävydestä suhteessa asetettuihin vaatimuksiin. Näiden tulosten perusteella pyritään myös selvittämään mahdollisia yhteneväisyyksiä aiempien arvosanojen suhteen jotta nähtäisiin mahdollinen keino arvioida opiskelijan taitotasoa ilman testausta.</p> <p>Valtaosa opiskelijoista ns. ”läpäisee” diagnostisen testin, eli omaa vaadittavan taitotason joka on edellytyksenä valmistuvalle ammattikorkeakouluopiskelijalle. Opiskelijoiden diagnostisesta testistä saaman tuloksen yhteneväisyys aiempiin arvosanoihin on myös oletetun kaltainen, jonka perusteella voidaan sanoa että karkeaa jaottelua aiempien arvosanojen perusteella voidaan perustellusti tehdä, mutta kovinkaan hienojakoista arviointia ei pystytä tekemään, ainakaan yksittäisen arvosanan perusteella. Arvosanojen yhdistelmät, päättötodistuksen arvosana yhdistetynä YO-kokeen arvosanaan, tuottavat kuitenkin jonkin verran laadukkaaman kuvan opiskelijan tasosta. Selkeä joukko opiskelijoista ei kuitenkaan ole arvioitavissa pelkkien aiempien arvosanojen perusteella, jolloin diagnostisen testin kyky arvioida ko. opiskelijan todellista tasoa on tarpeen.</p>	
<b>Avainsanat:</b> kielten opiskelu, diagnostinen testaus	



## TABLE OF CONTENTS

PREFACE

ABSTRACT

TIIVISTELMÄ

## TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
<b>2</b>	<b>LANGUAGE LEVEL AND LANGUAGE REQUIREMENTS AT UAS'S</b>	<b>4</b>
<b>3</b>	<b>DIAGNOSTIC TESTING AND THE METROPOLIA LANGUAGE TEST</b>	<b>6</b>
3.1	<b>The Metropolia Diagnostic Language Test</b>	<b>6</b>
3.2	<b>Other Diagnostic Tests</b>	<b>8</b>
3.2.1	<i>YKI (Yleiset kielitutkinnot – General Language Examinations)</i>	<i>8</i>
3.2.2	<i>DIALANG</i>	<i>9</i>
3.3	<b>Synopsis</b>	<b>10</b>
<b>4</b>	<b>METHOD AND MATERIAL</b>	<b>12</b>
4.1	<b>Background Data</b>	<b>13</b>
4.2	<b>Validity and Reliability</b>	<b>14</b>
<b>5</b>	<b>RESULTS OF DIAGNOSTIC LANGUAGE TESTS</b>	<b>16</b>
5.1	<b>Key for reading the results</b>	<b>16</b>
5.2	<b>Longitudinal Analysis of 2005 – 2009</b>	<b>17</b>
5.2.1	<i>Matriculation Examination Background in Swedish and English</i>	<i>18</i>
5.2.2	<i>High School Final Grade Background in Swedish and English</i>	<i>21</i>
5.2.3	<i>Correlation of Test Results with High School Grades and Matriculation Examination Results Combined</i>	<i>24</i>
5.3	<b>Crossectional Analysis of 2008 - 2009</b>	<b>27</b>
<b>6</b>	<b>DISCUSSION AND CONCLUSIONS</b>	<b>32</b>
6.1	<b>General</b>	<b>32</b>
6.2	<b>Results</b>	<b>33</b>
6.3	<b>Recommendations, suggestions &amp; personal thoughts</b>	<b>35</b>

## 1 INTRODUCTION

This thesis project analyzes the level of language skills in English and Swedish of students of five Universities of Applied Sciences. The data involves results from Metropolia over four years and additionally the results of four other UAS's during one year. Metropolia as the largest university of applied sciences (later UAS) in Finland currently accepts close to 3000 new students annually. A majority of these students take a diagnostic language skill test in Swedish and English to determine their current skill level and the need of preparatory courses before they participate in the language education. The decree on universities of applied sciences education states that a graduating student must possess sufficient level in both national languages, as well as in at least one additional language to properly cope in his or her professional field (UAS Decree 2003). Only students excluded of these requirements are those who have received their lower level education somewhere else than Finland. The Common European Framework of Reference for Languages (later CEFR) divides language skill levels to six categories; A1-A2, B1-B2 and C1-C2, which roughly divide so that A-level user has the very basic skills, B-level user is competent user and the C-level user has above average skills and abilities using foreign languages (CEFR 2003). Sufficient skill levels for graduating students are B1 for Swedish and B2 for English. The Swedish level B1 is thought to correlate with the sufficient level of skills mentioned in the decree on polytechnic education. The English level B2 is a decision made by the Language Group of the Rectors' Conference of Finnish Universities of Applied Sciences.

The research question thus concentrates on the following dilemma: Do the UAS students possess the skill level required in foreign languages?

Additionally, the possibility of making the diagnostic testing obsolete if the students' skill levels could be accurately determined by their previous grades is explored. Already some UAS's consider diagnostic

testing unnecessary as they see that the previous grades are accurate enough to determine the skill level, and in Metropolia also students are even exempted from testing if their matriculation examination grade is in the top two, E or L. This study will look in depth the possibility of excluding other grades as well, either viewing them to be sufficiently skilled, or instruct them directly to the preparatory courses, without the need for a laborious diagnostic test. In this study I will look into the grades and look for not only singular grades, but also the combinations of matriculation exam grades combined with the final grade to see if within one matriculation exam grade there is variation.

In this study the data is sampled from a period of four years, from August of 2005 to May of 2009, of diagnostic tests in both languages to see whether or not students actually meet these requirements, or are more preparatory courses or other additional education needed to meet the legislative requirements. Additionally one specific year will be looked at in cross-reference comparison between four other UAS's in Finland to see possible variation on a national level. These four UAS's use the very same diagnostic test developed in Metropolia, which make the results are highly comparable.

The sample of Metropolia students is nearly 6800, and out of this 70% come from high school background and 30% come from various vocational schools. Since the majority of students come from high school, a closer look will be taken on the grades graduating high school students receive. In Finland student must participate in matriculation exam before graduating from high school, and will receive a result from said exam and also a final grade from the courses taken during high school. Students graduating from high school will therefore have two grades, one from matriculation exam grading in ascending order A, B, C, M, E and L, and another numerical grade in a scale from 4 to 10. In this study I will look in to these grades and see their relevance to the grades the students get in the diagnostic test. The results received from vocational school students was so incoherent that it came across

as unusable. In vocational schools two different scales are used, other ranging from 1 to 5 in ascending order, and other using the similar scale than in high schools, ranging from 4 to 10. These scales are used happily mixed, which made the use of these results nigh impossible. This led to the decision that this study would focus solely on high school students.

## 2 LANGUAGE LEVEL AND LANGUAGE REQUIREMENTS AT UAS'S

This section focuses on the requirements for language education and acceptable skill level for students of UAS's. Even though both languages, Swedish and English, are considered foreign languages in this study, their status in education differs greatly. Being bilingual country, in Finland also Swedish language is considered national language, and therefore enjoys different benefits in front of the law.

The law regulating language requirements for state personnel requires that a person in the public sector must have a certain level of understanding in both national languages, and this imperative is also included in the UAS decree stating that a person graduating from UAS must fulfil the language requirements for state personnel expressed in UAS certificates as follows:

The student has acquired the language skills in Swedish/Finnish required of state officials functioning in a position requiring a higher education in a bilingual office according to the Act (424/2003, 6 §) with a grade of good/satisfactory in written language skills and good/satisfactory in spoken language skills. The student has also acquired the language skills in Swedish/Finnish necessary for practicing the profession and for further professional development (Decree 352/2003, 8 §, 1).

The mission statement formulated by the Foreign Language Group of ARENE, Rectors' Conference of Finnish Universities of Applied Sciences, states that this level of skill in Swedish language is sufficient for student graduating from UAS, and that it correlates with CEFR-level B1, which is considered to be the level of requirement for UAS students.

As for other languages besides Swedish Legislation states, that the student must possess "such written and oral skills as are needed for the profession and professional development in the field." (Universities of Applied Sciences Decree 2003) This decree is detailed as a mission statement formulated by the Foreign Language Group of ARENE,

Rectors' Conference of Finnish Universities of Applied Sciences as follows:

The mission of universities of applied sciences (UASs) is to provide the student with the language and communications skills needed in specialist positions in the increasingly international working environment and business life.

The UAS language education emphasizes students' capabilities for interaction across languages and cultures and the student's own professional development. The development of Language and Communication for Professional Purposes is based on national and international research and aims to sustain international transparency and competitiveness. (ARENE language group 2006)

The Foreign Language Group of ARENE states that the level of understanding in foreign language should be higher than in Swedish, and agreed that the level should be a B2 on CEFR-scale.

### **3 DIAGNOSTIC TESTING AND THE METROPOLIA LANGUAGE TEST**

In this section a more detailed look will be taken at the diagnostic test in Metropolia UAS, the target test of this study. Two other corresponding tests will also be introduced, EU-funded, multi-national DIALANG, and YKI, National Foreign Language Certificate of Language Proficiency, organized by Finnish National Board of Education. Of these tests only the DIALANG test is available to the general public free of charge. The YKI is governed by the Board of Education, and is held only twice a year and costs up to a 100 euros depending on the level of the test taken, whereas the Metropolia test is used only in Metropolia UAS and six other UAS's that have obtained a license to use said test in their own student evaluation process.

#### **3.1 The Metropolia Diagnostic Language Test**

Metropolia UAS uses a diagnostic test specifically made for Metropolia by Metropolia students. There are four different, field-specific variations of the test, providing opportunity to use the test in different fields, relevant to students. It is designed to be focusing more on work related situations rather than general language skills. The test is web-based java-application, so it is both easy to take and easy to manage or alter. Its database also provides the organisers ability to access students via email and retrieve results quite easily in one excel-document. The test is designed to be an assessment of language skills and it will provide the participant with immediate result on the CEFR scales (CEFR 2001). The test is fairly quick to complete, lasting only for 90 minutes, yet being quite diverse testing several aspects of language use. Test consists of six different parts, all of which are work oriented situations measuring the language skills in participants own field of education. There are four content specific tests for culture and arts, social services and health care, technology and services. There are also multiple tests in each category, giving the organizers the opportunity to alter the tests between groups to minimize the opportuni-

ties of cheating. The languages available languages are English and Swedish, taught as obligatory languages at UAS's.

The test is composed of six parts, named alphabetically from A to F. The first part is a short text after which ten true or false statements are made, six in Finnish and four in English. This part tests both *reading comprehension* and *understanding of vocabulary and structures*. The second part is ten multiple choice questions, in which students must choose the most polite answer or reply to a phrase given. This part tests mainly *conversation skills* and understanding of politeness, but also grammar. The third part is a cloze test, with a text with every sixth or so word missing, and students must fill in the blanks. Based on testing research the cloze test is the most reliable format to test overall language skills. There are 30 words missing from the text. This part tests *general linguistic competence*, familiarity with dialogue, *politeness*, *vocabulary and structures*. The D section is a multiple choice question again, asking for the correct translation to 15 Finnish words. This part tests *vocabulary* and structural analysis. The final part is filling out the missing words, with prompts given in Finnish. 30 *grammatical items* are missing, and participant must fill them out, focusing on tenses and conjugations of the words. In the last part students must choose a correct option from the list to provide the best question for the five answers given. All these parts add up to a maximum score of 100. Table x summarizes the content of the diagnostic test.

Name of the Section	Subject	Number of questions	Recommended time used
Background	Student fills out background information		5 min.
A / Short text	Text comprehension, also language.	<b>10</b>	15 min.
B / Conversation	General conversational skills, understanding of politeness. Vocabulary, grammar.	<b>10</b>	10 min.
C / Cloze test	General linguistic competence, fa-	<b>30</b>	20 min.

	miliarity with dialogue, politeness, vocabulary, structures.		
D / Vocabulary	Vocabulary and structures	<b>15</b>	10 min.
E / Slot filling	Knowledge of verb structures, tenses, nouns, adjectives.	<b>30</b>	20 min.
F / Question formulating	Grammatical structures.	<b>5</b>	5 min.

*Table X. Components of the Metropolia diagnostic language test.*

Table x shows how the test covers many aspects of language proficiency, such as reading comprehension, dialogue and politeness discourse, vocabulary, grammar and text production.. There are, however, aspects the test does not measure. The main aspects not being tested in this computer-based test are *oral language skills*, and *creative writing skills*, both of which are fairly difficult to test in a computer-based test providing immediate results. Both of these would require a human interaction in assessing results.

## 3.2 Other Diagnostic Tests

Two other language assessment tests have been studied for additional perspective on language testing in Finland. First one is The National Foreign Language

### 3.2.1 YKI (Yleiset kielitutkinnot – General Language Examinations)

The National Foreign Language Certificate of Language Proficiency is a language test issued by the University of Jyväskylä and sanctioned by the Ministry of Education in Finland. The test is offered in nine languages, English, Spanish, Italian, French, Swedish, Sámi, German, Finnish and Russian. The test is divided in to three levels, lower, middle and upper levels, which are all divided in to two sub levels correlating with six levels of the CEFR-levels. Participants must themselves decide which level they want to take the test in. The test is open to everyone, and it costs between 65 and 140 euros per test, depending

on the level of the test, the basic level being the cheapest. These tests are held only twice a year, except the Finnish test which is held four times. The test is held at 72 different locations spread over Finland, usually in educational establishments. The tests are quite lengthy, ranging from just over two hours for the middle level test to over four hours for the advanced level test.

The tests are divided into segments, each segment taking between 30 and 60 minutes of time. If a participant were to fail any of the segments in the test, the entire test would fail, and the whole test needed to be retaken for a passing grade. All tests include the following four segments; Writing, Sentence structure and vocabulary, Speech comprehension and Speaking. The basic level also includes Reading Comprehension, and the advanced includes Reading Comprehension and an interview. The main part the test is missing is creative writing. In its vocabulary and general theme the test is very universal, and thus doesn't provide the necessary information about the understanding of professional language often needed in different universities and UAS's. Also problematic is the demand for the participant to beforehand determine the tested level, Beginner, Intermediate or Advanced, and then he or she can only be scored within this level. For example user passing a test in Beginner level could possibly be an Intermediate user lacking confidence to enrol to a more appropriate level, and getting a high grade in Beginner test doesn't provide the participants with an accurate result of one's skills.

### **3.2.2 DIALANG**

Dialang is a joint European project which was initiated with the purpose to develop a diagnostic language assessment tool for EU countries. It was developed between the years of 1996 and 2004, and it was funded by the European Union's Socrates Lingua project. Projects' first phase was coordinated by University of Jyväskylä's Centre for Applied Language Studies, and the second and third phase by the

Freie Universität Berlin, who are still mainly responsible for project coordination. The outcome of this development project was a free of charge language diagnostic tool for all European citizens. Dialang is not intended to be used for certification purposes, but rather as a self-assessment tool for personal use and to provide language users and learners with diagnostic information about their language skills independent of any formal education system. The test is available in all official EU languages as well as Irish, Icelandic and Norwegian, in total of 14 languages, and will cover all skill levels from beginners to advanced. Once finished the test will provide the user with a result according to the CEFR-scale. Dialang offers separate modules for reading, writing, listening, grammatical structures and vocabulary. The modules incorporate self-assessment, externally assessed tasks, and feedback. The clients decide which modules they want to take, in which order, and whether they want to take one or several modules at one go.

Because of the immediate CEFR-result that is provided to the participant, the Dialang test is used in numerous educational institutes for students' self-assessment, but also for diagnostic purposes assessing the incoming students. The main problem with this sort of "abuse" of the test is that it is very generalized in its vocabulary and theme and the language used in many universities and UAS's is very content specific, and therefore the skill level acquired in this generalized test may not be sufficient for the possible studies ahead. Also the easy access to the test enables the students to download the test for themselves and practice at home achieving a better result through short-sighted practicing in hope of avoiding the possible additional courses allocated for students with poor results.

### **3.3 Synopsis**

These tests are available to determine the language skill level of student entering an UAS. This diagnostic work is needed to identify the students that may require some special attention during their studies

at UAS in order to graduate with sufficient skills of foreign languages. Out of these three tests only Dialang is openly available free of charge, and the YKI for some charge, but the Metropolia test is only available for the students of UAS's that use said test. The Metropolia test is also only one with any field-oriented material providing different tests with unique question sets for different branches of education. The other two are more general in their test settings. Only YKI measures creative writing, mainly probably because it is the most difficult type to grade, and also YKI is the only one with an interview measuring the participants' oral skills. This is mainly due to the nature of these tests; Metropolia and Dialang are both electronic web-based tests and these two types of tests require human interaction either grading or receiving the test.

#### 4 METHOD AND MATERIAL

This study was conducted as a quantitative statistical analysis, using the existing raw data from tests taken between August of 2005 and May 2009. All the four content specific tests have multiple versions and both languages combined there are a total of 23 different tests available. For this study the sample had to be narrowed down, for the amount to be reasonable, so the tests most often used were selected for a largest possible data. Tests number 1 and 2 from both English and Swedish were selected, and from three different branches, Health Care, Culture and Arts and Technology. This gave a total of six tests from both languages to work with, and a sample of over 1500 students per year in both languages. The timeframe consists of four school years; 05-06, 06-07, 07-08 and 08-09, total timeline ranging from 1<sup>st</sup> of August 2005 to 31<sup>st</sup> of May 2009.

ENGLISH	05-06	06-07	07-08	08-09
	<b>1965</b>	<b>1042</b>	<b>1281</b>	<b>1287</b>
SWEDISH	05-06	06-07	07-08	08-09
	<b>1903</b>	<b>1522</b>	<b>1686</b>	<b>1672</b>

*Table 1: Yearly Sample sizes from Metropolia*

Besides this longitudinal study, for reference a cross-sectional study was conducted where the results of four UASs were set aside with the results of Metropolia for one year, from August 2008 to May 2009. For this part results of the diagnostic tests were asked from four other UAS's that use the Metropolia test giving comparable results from their diagnostic tests. The purpose of this comparison is to see whether or not the results from Metropolia are comparable nationwide. The other Universities of Applied Sciences are Savonia UAS, Jyväskylä UAS, Tampere UAS and Rovaniemi AMK, but in this study the universities will be discussed anonymously, only describing them as Specimen UAS 1 through 4, numbered in random order.

## 4.1 Background Data

The sample of Metropolia students is 6783 in Swedish, and 5575 in English, and out of this 70% come from high school background and 30% come from various vocational schools. Since the majority of students come from high school, a closer look will be taken on the grades graduating high school students receive. For the cross-sectional part focusing on the school year 2008-09, a total sample of 1303 for the Swedish and 1134 for the English part were collected from the four participating UAS's and compared these results with correlating results from Metropolia from that year. The Metropolia samples for this year were 1672 for Swedish and 1287 for English. The smaller samples all round for English language can be explained with the conditions for the sample collected, in both language tests there are several different tests that can be used, and for both languages only first two tests were chosen. In English there are more tests to choose from which explains the scattering of students more widely.

From the table 1 can be seen that the yearly variation between samplesizes is quite reasonable if the first ever year 2005-06 is excluded. The diagnostic test was first introduced in 2005, and on the first year also some students from previous years took the test and thus appear on the statistics.

For the cross-sectional study conducted, the sample from four other UAS's was collected, and theirs samples put together add up to almost equal amount as Metropolia's sample alone. This considerably smaller sample size can be explained not only with the considerable difference in sizes between the schools, but also the differentiation in

branches. Because from Metropolia only three branches were selected, Healthcare, Culture and Arts and Technology, the intention was to collect the sample from the other UAS's from roughly similar branches if possible. These branches vary in different schools, and all branches were not even present in all of the schools.

## 4.2 Validity and Reliability

Validity represents how well a variable measures what it is supposed to. Validity is important in descriptive studies: if the validity of the main variables is poor, thousands rather than hundreds of subjects are needed. Reliability tells you how reproducible your measures are on a retest, so it impacts experimental studies: the more reliable a measure, the fewer subjects you need to see a small change in the measure.

	<b>Validity</b>	<b>Reliability</b>
<b>Internal</b>	Is there a causal relationship between the treatment and the outcomes?	Are the results internally consistent among subjects?
<b>External</b>	Can we generalize these results to other settings?	Are the results consistent across uses, such as the use of a test to measure achievement?
	<b>Construct Validity</b> How do the concepts we use to quantify our results actually relate to the corresponding treatment or measurement? (e.g., "engagement" to "measurement of engagement")	<b>Inter-rater Reliability</b> (a form of external reliability) How consistent are the observational ratings among a group of observers?
	<b>Conclusion Validity</b> Is there a relationship between the treatment and our conclusions? (e.g., "novel approach" to "increased engagement")	

**Table 2:** *The structure of validity and reliability*

Following the structure pictured above this section will go through the reliability and validity of this study. Regarding the internal validity this study the results follow in general the presumptions made beforehand regarding the students' grades and the correlation of them with the results from the diagnostic test. Although the test deepened the understanding of the correlation and refined the level of correlation between pre-existing grades and test results, the general idea of students with better backgrounds would do better in diagnostic testing as well stood.

If looked at the generalizability of the results, it is hard to see how any other study would result in dissimilar results. The sample size in this study is relatively large, and it consists of several years of students entering Metropolia, making the sample credible. Conclusion validity represents the degree to which conclusions reached about relationships between variables are justified, and in this study the relationships between the variables seem to stand quite well, as they are in line throughout the study, and also between languages. The internal reliability of the study is relatively good, as the results are similar not only through years in Metropolia's case, but also through different UAS's as can be seen in the cross-sectional study. This would implement that the data is not only reproducible not only within Metropolia and this study, but also with other similar studies measuring the level of language skills in UAS students, and the correlation between high school grades and the results of the diagnostic test.

## 5 RESULTS OF DIAGNOSTIC LANGUAGE TESTS

### 5.1 Key for reading the results

In Finnish Universities of Applied Sciences there are numerous different fields of education, of which in this study only three were looked in to, Technology, Social studies and health care, and Culture. This study uses Finnish based abbreviations “Sote” and “Kul” of Social studies and health care and Culture, accordingly. Also the abbreviation of “Tech” will be used when discussed about the Technology. Common abbreviation used is the UAS meaning a University of Applied Sciences.

A little background information about the Finnish Matriculation Examination might also be useful. The Exam is nationwide test held bi-annually, in spring and fall, in all Finnish high schools at the same time. The Matriculation Examination was first arranged in Finland in 1852. In the beginning, the examination was the entrance examination to Helsinki University, and in it one had to show sufficient evidence of an all-round education and a knowledge of Latin.

Nowadays, the purpose of the examination is to discover whether pupils have assimilated the knowledge and skills required by the curriculum for the upper secondary school and whether they have reached an adequate level of maturity in line with that school’s goals. Passing the Matriculation Examination entitles the candidate to continue his or her studies at university. (Matriculation exam board)

The Matriculation Examination is regulated by Section 18 (766/2004) of the Upper Secondary School Act, the Act on the Organization of the Matriculation Examination(672/2005) and the Government Decree on the Matriculation Examination (915/2005).

The Matriculation Examination Board is responsible for administering the examination, its arrangements and execution. The Board issues guidelines on the contents, the arrangements and the assessment of the tests.

In Finnish education the other option for secondary education other than High School is Vocational school. In Finland 42% of students enter Vocational School after basic education, and 50% go to High school (stat.fi)

## **5.2 Longitudinal Analysis of 2005 – 2009**

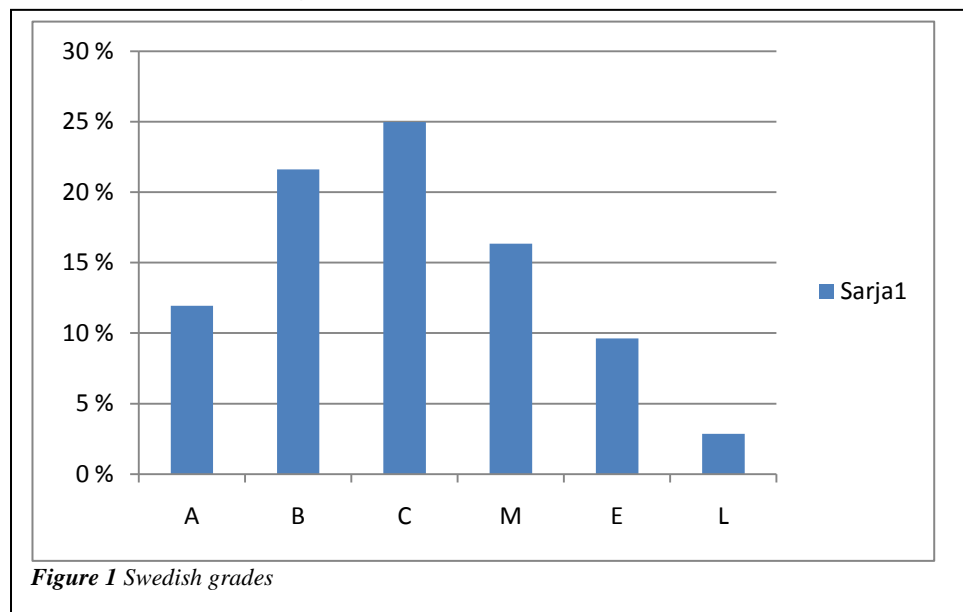
This section will focus on the results of the study, going through the end results from the two languages in Metropolia during the four year span. The results are examined in the light of both research questions: do the students possess the skills required, but also do the pre-existing grades tell the current skill level with acceptable accuracy. The previous results will be looked at in three categories, only Matriculation Exam results, only High School final grade, and the two grades combined to see if and what grades are most informative of the students skill level.

Exactly 70% of students entering Metropolia have graduated from high school prior to entering and 30% have various forms of vocational school backgrounds. Unfortunately the data collected from the vocational school students is not reliable enough to do a comparative analysis in the way the high school results will be analyzed. This is mainly to do with the fact that the grades students supposedly are given in vocational schools range either from 1 through 5, or like in high schools from 4 through 10. Unfortunately the data does not differentiate these different styles of marking, therefore giving results ranging from 1 through 10 making it virtually impossible to compare these results as it's impossible to determine the scale used. Thus the real focus in this study will be on the High School graduates which make for the vast majority of students entering Metropolia.

### 5.2.1 Matriculation Examination Background in Swedish and English

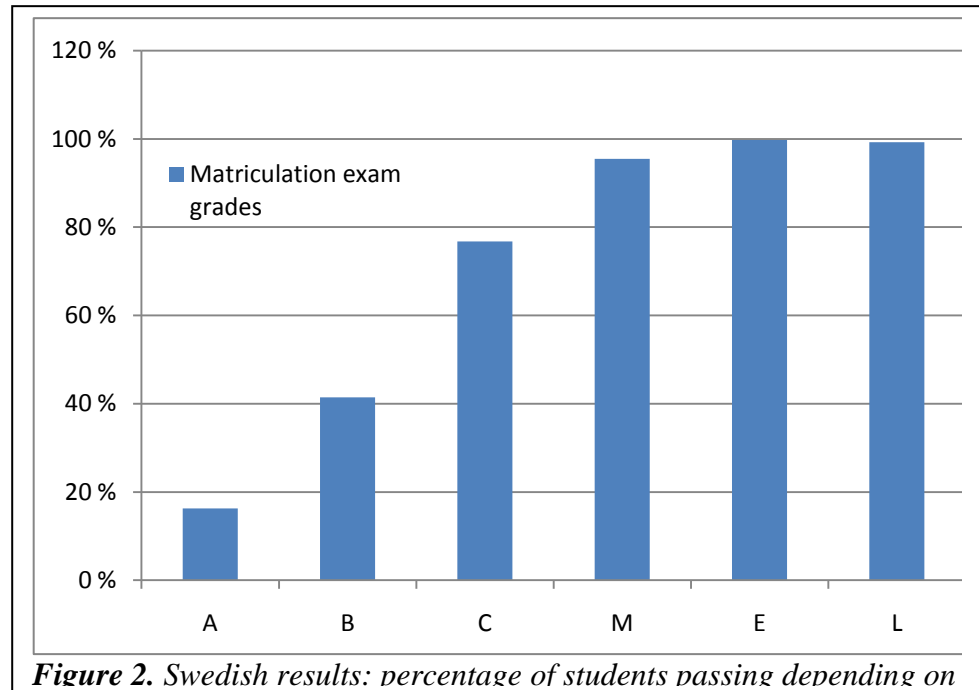
The high school students wishing to graduate from high school must take part in Matriculation Exam which is nationwide tests which are graded centralized by the Matriculation Examination Board. The grading for the Matriculation Exam follows the normal distribution, and is divided into seven verbal grades, L, E, M, C, B and A as the lowest passing grade. The highest grade is L (laudatur) and it is given to 5% of the students. The following grades are divided so that 15% get E, 20% get M, 24% get C, 20% get B and 11% get A. The

As the study was done from two languages, Swedish and English, the results are also processed apart; first the Swedish results and then the English results. The distribution of the grades follows the normal distribution, with slight emphasis on the top end of the grades as can be seen in the first, graph 1.



There is also a group of around 10% of students who have not entered their matriculation exam grade. When looking the group of students passing the test, the students with higher grades, like L and E, are quite confidently passing the test, and that the lower grades such as the A and B are failing the test in numbers. The problematic part is the middle ground, M and especially C grades, which are doing quite alright, majority of the students passing, but not so much that any

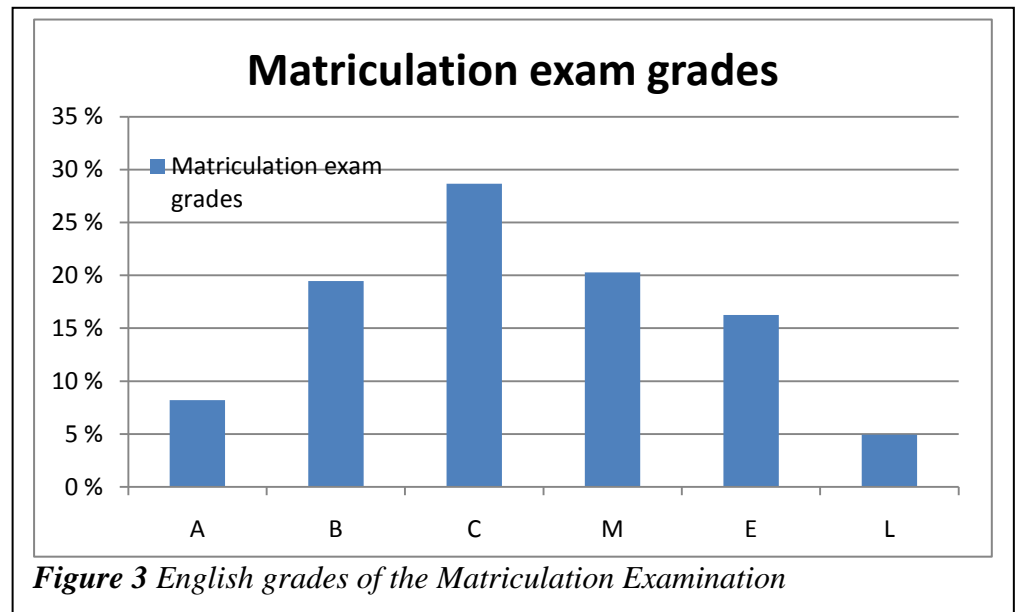
rules could be made. This problem is addressed later when the combined grades are looked into. In the figure 2 we see the level of which the matriculation exam result grades have passed the test, e.g. 100% of students with L grade have passed the test.



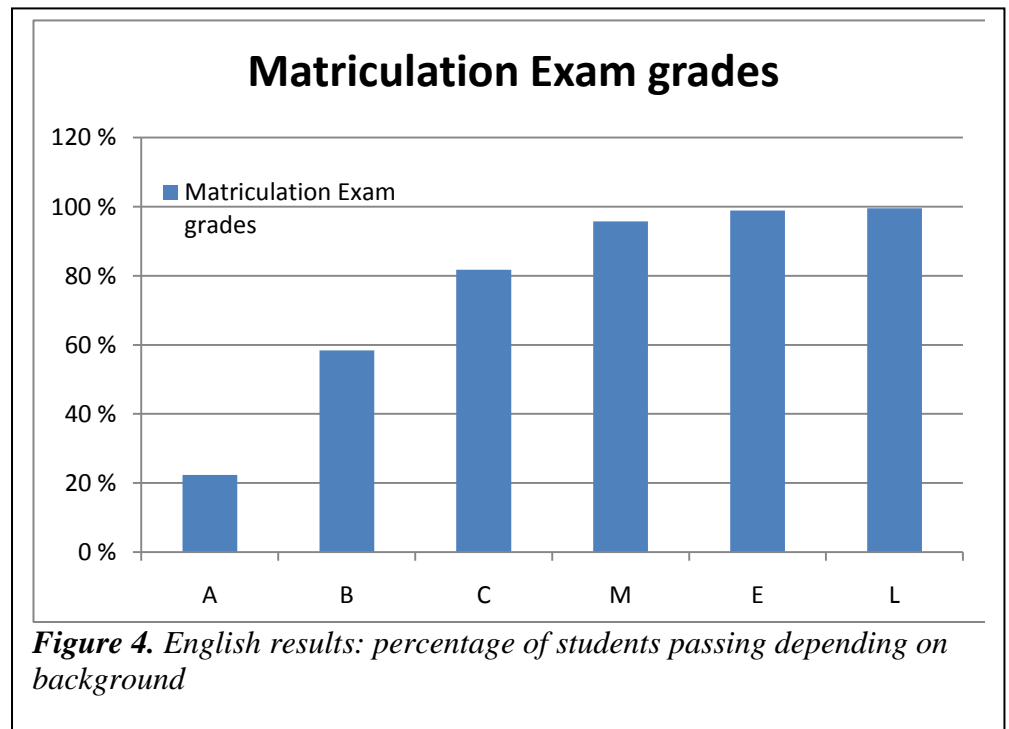
Students with a grade of M or higher can be seen to pass the test with quite the consistency, whereas of the students with A only one in five possess the desired skill level in Swedish.

The total number of students with matriculation exam grade is 4740 in Swedish test, and the number who passed the test is 2836, giving a total percentage of passing grades of **59,83%**.

In English there was total number of 4323 students with High School background, approximately 68% of students, and the grades of the matriculation exam are pretty alike to those of Swedish test, if only slightly higher. This correlates with the general presumption that students have better skills in English than in other foreign languages.



As can be seen in the figure 4, the results are as per the normal distribution, almost one third of the students entering Metropolia have a grade C, similar amount has a grade A or B, and the rest is spread over the higher grades. In this stat there was only a 2% group which did not declare a grade so it is more reliable than in Swedish. When looking at the results in the light of passing the test, we can see that the overall skill level is higher, and over 80% of the students with grade C from Matriculation Exam passed the test.

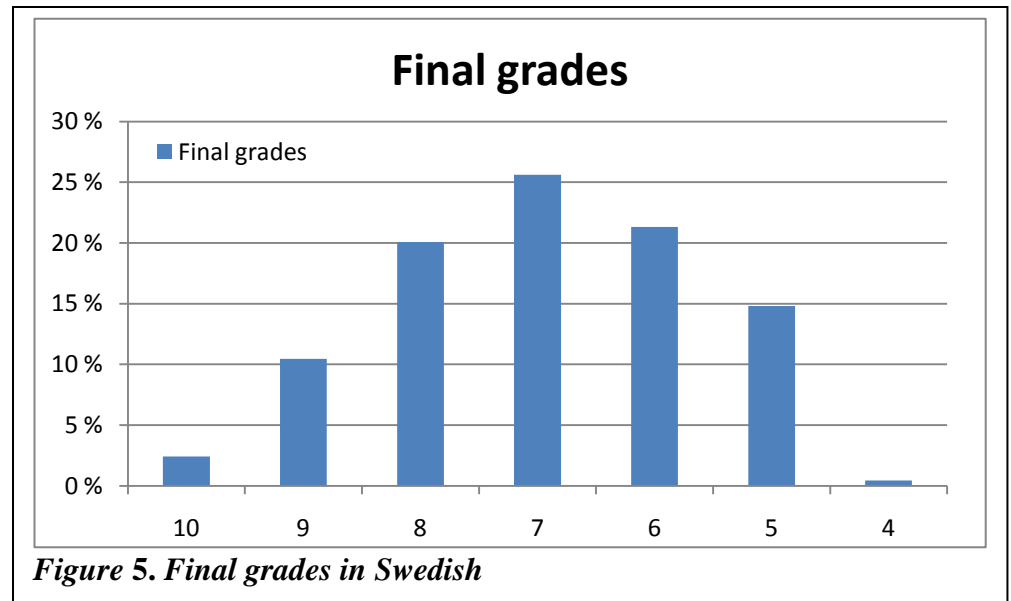


The level of skills is better than in Swedish, and a total number of 3367 students passed the test out of 4323, giving a total percentage of passing students **77,89%**

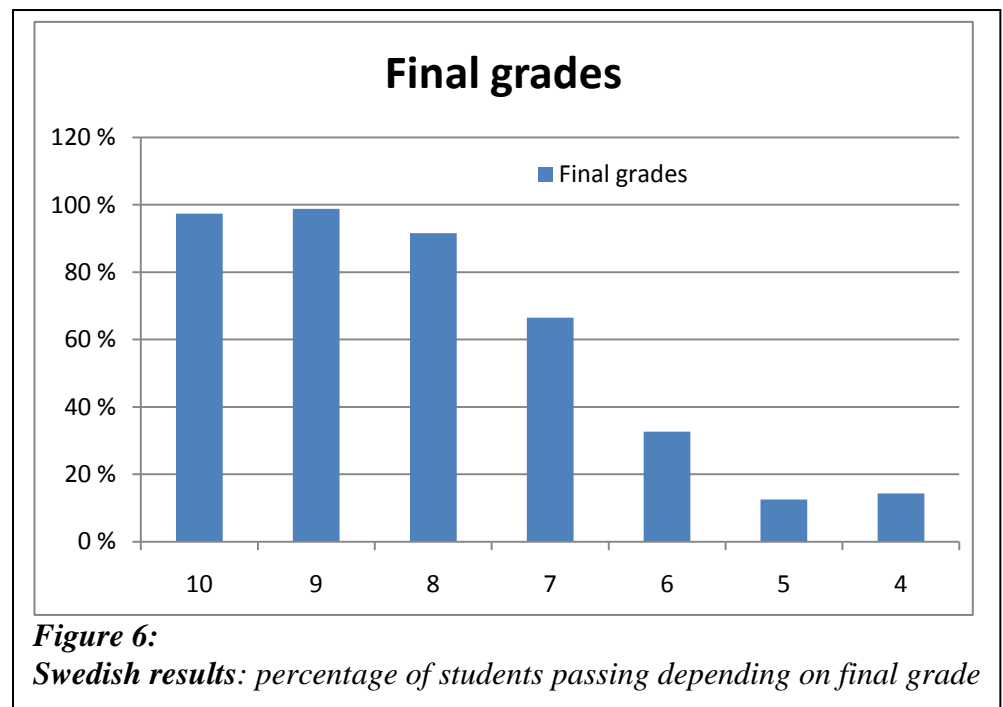
### **5.2.2 High School Final Grade Background in Swedish and English**

Students graduating from High School also get a certificate that lists all the final grades of the student. These grades are a mean value of the grades of all the courses the student has passed in one subject during the three years of High School. It can be argued that this grade, consisting of multiple courses taken spread over several years, would actually give more accurate information regarding the students' actual level of skills. Unfortunately the grades are not comparable against each other as the students have the freedom to choose the courses they attend in High School, and they vary from very basic courses to very advanced and specified writing courses. A high grade from an easy course does not require the same amount of skill as a mediocre grade from more advanced course. Also the number of courses varies between students, one might only take the bare minimum required, whereas another goes through all the courses available. Also the courses are graded by the teacher responsible of said course, and the grades are only a subjective opinion of the teacher of the students' performance in that course.

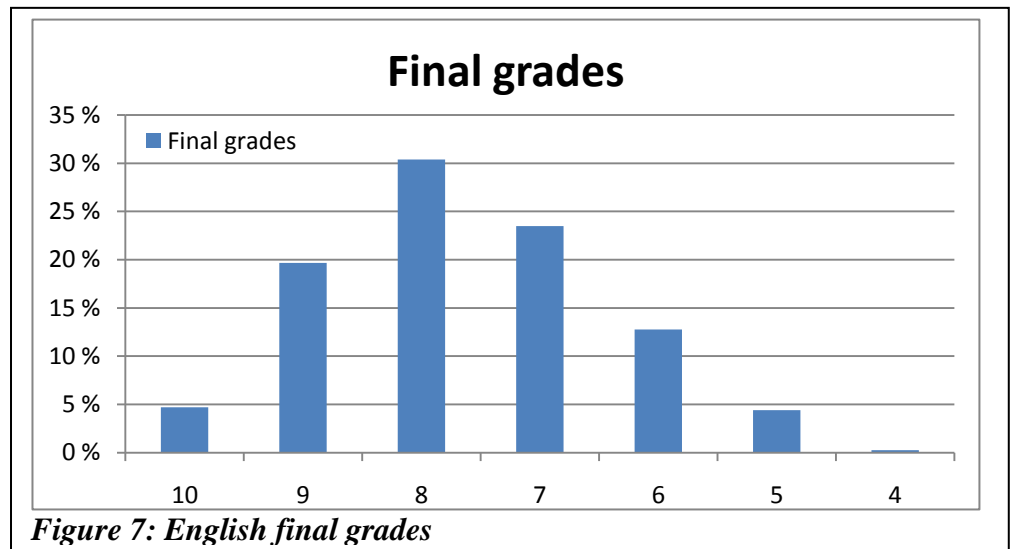
This section will go through the final grades in both languages, Swedish and English. In Finnish High School the scale goes from 4 to 10, 4 being basically failed, and 5 being the lowest passing grade. The first figure, figure 5, shows the distribution of the grades of the students in Swedish.



In Swedish the total data regarding final grades was 4740, but unlike in Matriculation Exam grades, only around 5% of students did not disclose their final grade. When looking at the level of students passing the test a different sort of results can be seen than in matriculation examination results. None of the groups cleared the test with 100% certainty, but of the three highest grades over 90% did pass the test, and then it quickly plummets so that out of students that received grade 6 less than 40% passed.

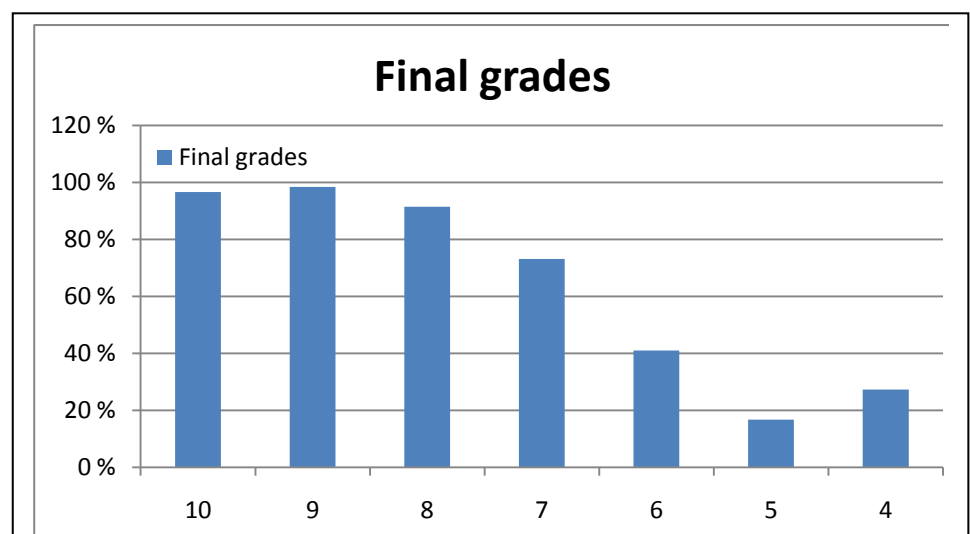


In English the overall level is much higher, and far more students have gotten a higher grade, 55% of the students have gotten a grade within the three highest grades, but this doesn't affect the level of students within grades passing the test, that figure is very much alike to that in Swedish, even though the background results showed improvement compared to Swedish. Total data of 4323 was available from the English test, and much like in Matriculation Exam grades there was a small fraction of less than 5% that did not enter their final grades.



**Figure 7: English final grades**

In figure 7 it can be seen clearly how the emphasis on the grades has moved towards the front and the higher grades, but in figure 8 it is obvious that this doesn't realize into that much better level of passing grades. Grade 6 still struggles to get over 40% of students passing, and lower than that basically one in five passes.



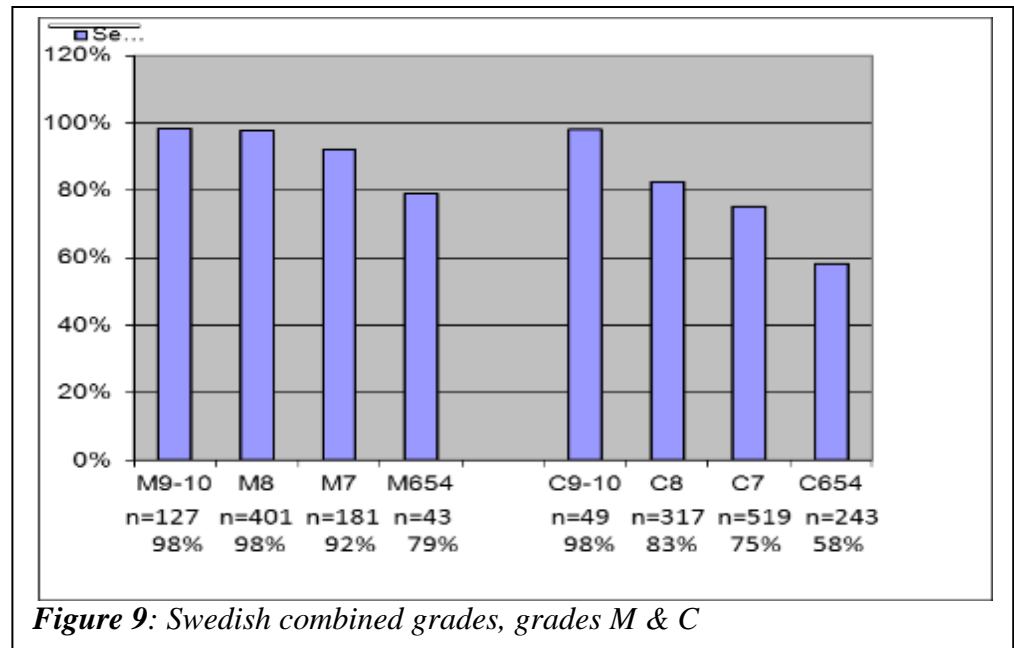
**Figure 8:**  
**English results: percentage of students passing depending on final grade**

Very much like in Matriculation Examination grades, students that have gotten a high grade as their final grade did well also in the diagnostic test. Problematic issue is, again much like in Matriculation Examination results, the hazy middle ground that is quite divided, many of the students do get a required result from the diagnostic test, but a large group is also lacking in skill which denies any possibility to make any all-inclusive decisions regarding this group. Therefore the third part will try to look into the combination of Matriculation Examination grades and final grades to see if there might be any way to differentiate the students with sufficient skill level and those in need of more preparatory education.

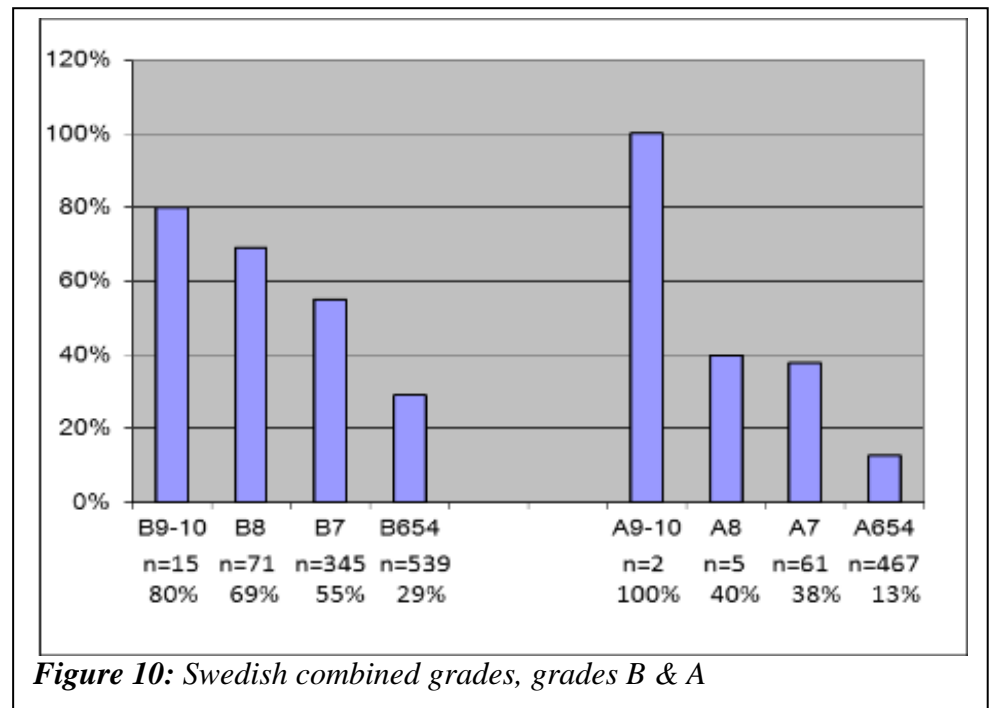
### **5.2.3 Correlation of Test Results with High School Grades and Matriculation Examination Results Combined**

After looking at the Matriculation Examination grades and the final grades separately, quickly an idea arose to combine these two grades in an effort to try and see if better analysis could be made of the students in the undecided middle-ground grades. It was obvious that students that have gotten a high grade from Matriculation Exam needed not to be looked into, as they would pass the test more often than not, so Matriculation Exam grades L and E were overlooked, and focus was shifted to the remaining grades, M, C, B and A.

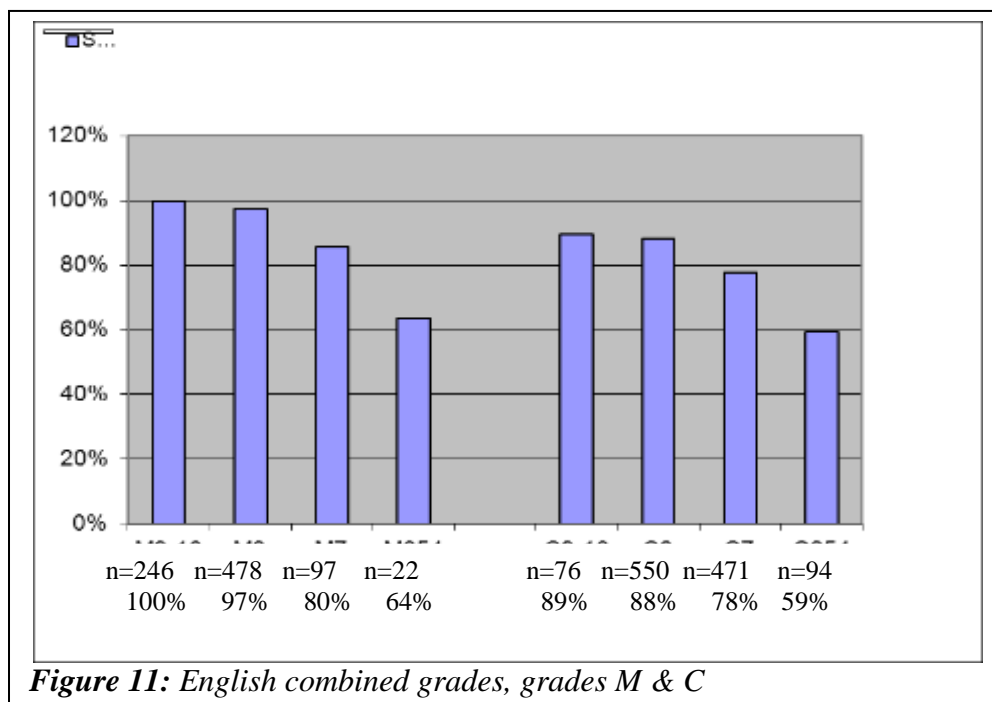
When looking at the combined grades in both languages it is obvious that the level of skills varies between students within one Matriculation Exam grade depending on the final grade it is accompanied with. It can be seen clearly in figure 8, when looking at the students with a grade C; the ones with 10 or 9 as their final grade are actually closer to the level of students with a grade M, whereas the students with lower numbers as their final grade are actually closer to a level of grade B. The gap between grades within C is over 40 percent.



In lower grades A&B the amount of data makes a slight problem, as the reliability suffers from much lower amounts of students, less than ten students compared to several hundred in previous grades. These smaller amounts combined with the fact that the background data is unverified and comes straight from the students themselves, makes the possibility of false data is very real. These sort of comedic answers easily mess up the data in such small samples.

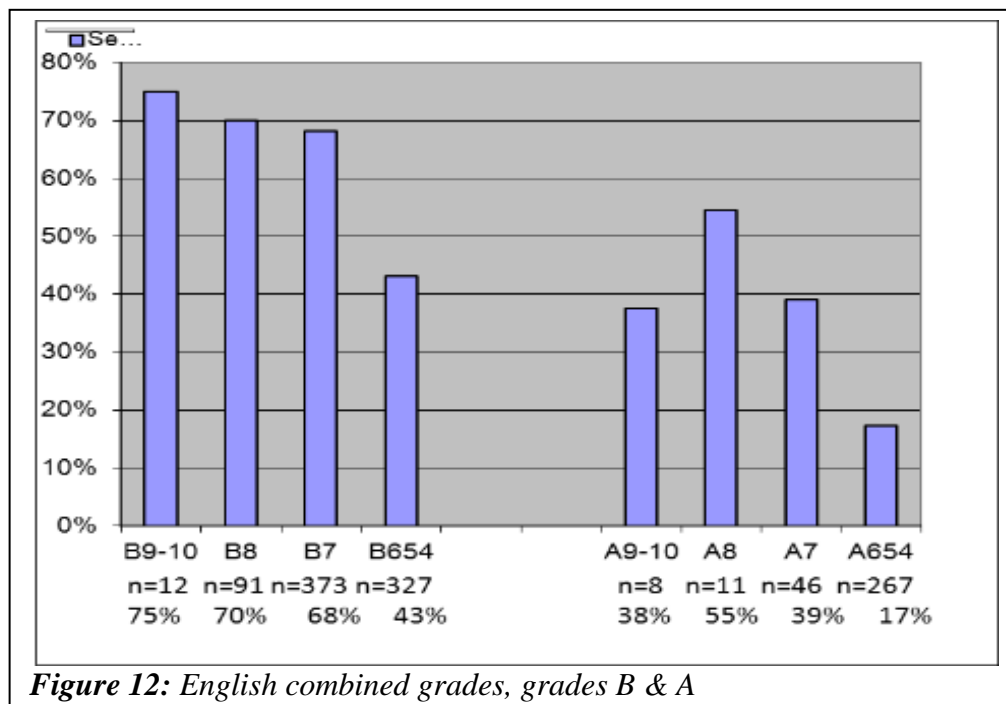


In English results, found in figures 10 and 11 an improvement very similar to previous English figures can be seen, but within a single grade there still are somewhat large variations. For example in grade M the higher grades 9 and 10 are very close to a perfect 100% passing rate, whereas lower grades like 6 and 5 are at only slightly over 60%



**Figure 11:** English combined grades, grades M & C

Again, much like in Swedish results, a problem arises with the small amount of results. When comparing results with several hundreds of students against those with only a handful of students the validity of the comparison suffers. The lower matriculation exam grades prove to be very diverse like they were in Swedish, within grade B the gap is well over 30% between students with higher final grades and those with lower grades.



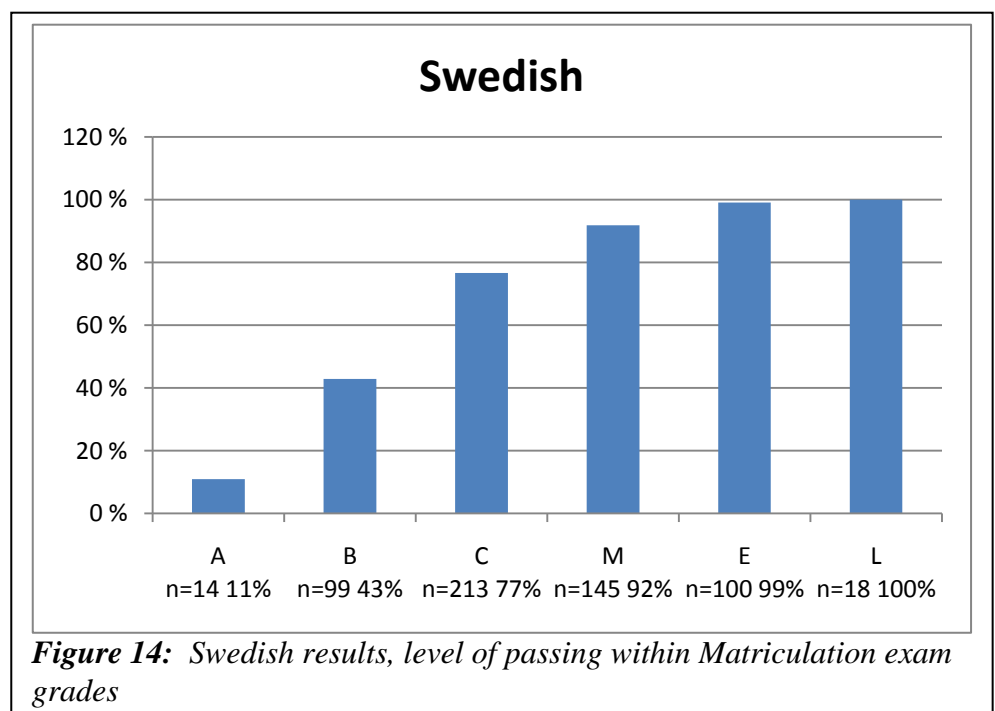
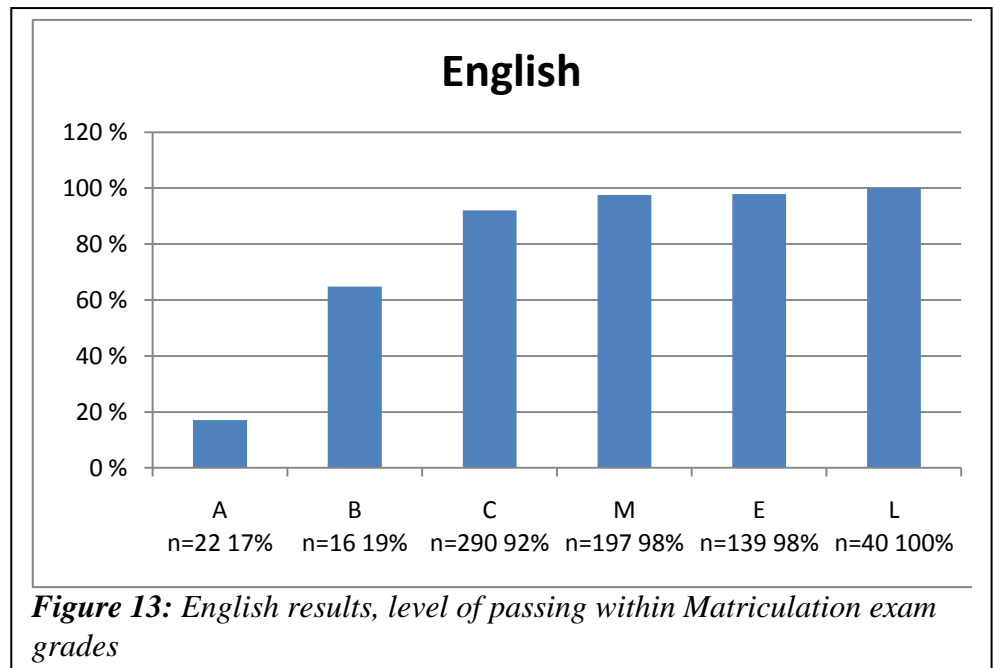
As a conclusion it can be said that a level of skills varies quite a bit between students inside one matriculation exam grade, and therefore judging any of these mid-to-lower grades as a uniform group would be unwise. Also a quite predictable decline in final grades can be seen as the matriculation exam grades come down, more students got a six or less as their final grade within grade A than in grades M and C combined.

### 5.3 Crossectional Analysis of 2008 - 2009

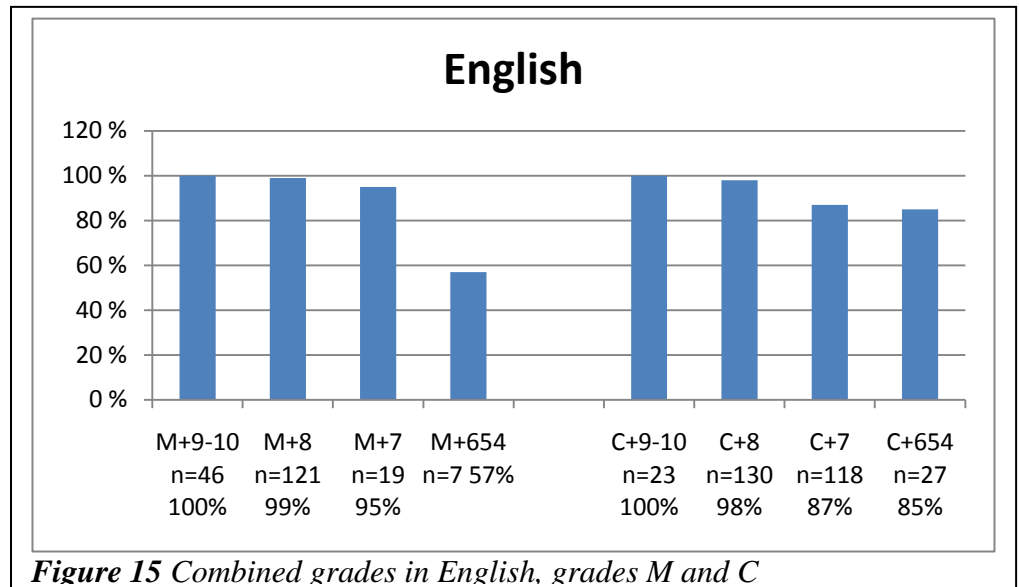
When putting together the data from Metropolia a question arose concerning the validity of the results, as they would be collected only from one UAS and thus the results would only represent Metropolia, and not UAS's in general. To widen the research, a crossectional analysis was conducted, and this focused only on one year, 2008-09, and included four other UAS's addition to Metropolia. JAMK University of Applied Sciences from Jyväskylä, Savonia UAS from Kuopio, Tampere University of Applied Sciences and RAMK UAS from Rovaniemi all use the Metropolia Diagnostic test to assess their incoming students, so the data collected was comparable to the data from Metro-

polia. These UAS's were asked to give results from similar or close to similar fields of study, and they provided a sufficient amount, a combined data of 1300 Swedish results and 1200 in English. As the purpose was not to rate the schools or to put them in order, this combined material was analysed as one, and the results were collected as they would be one UAS in comparison with Metropolia.

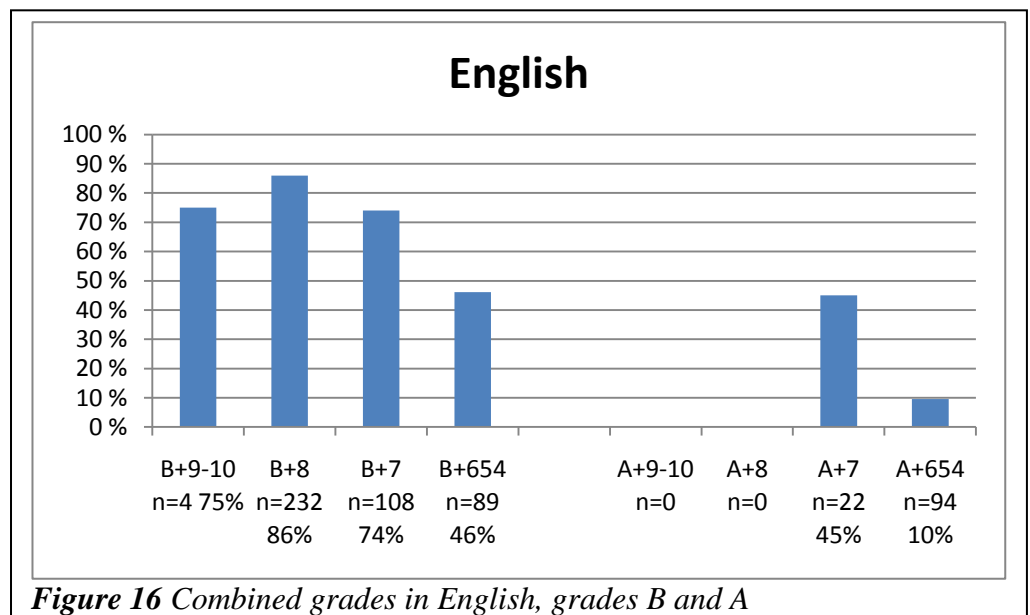
The results were most interesting, even though the UAS's are situated far apart across Finland, the results are very close to those collected from Metropolia.



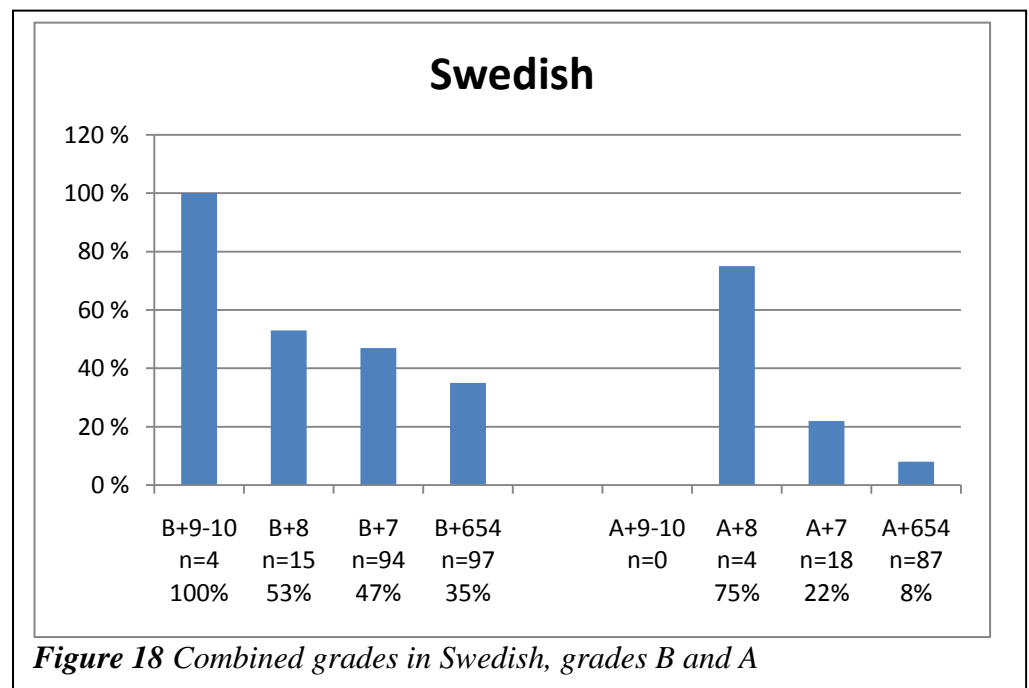
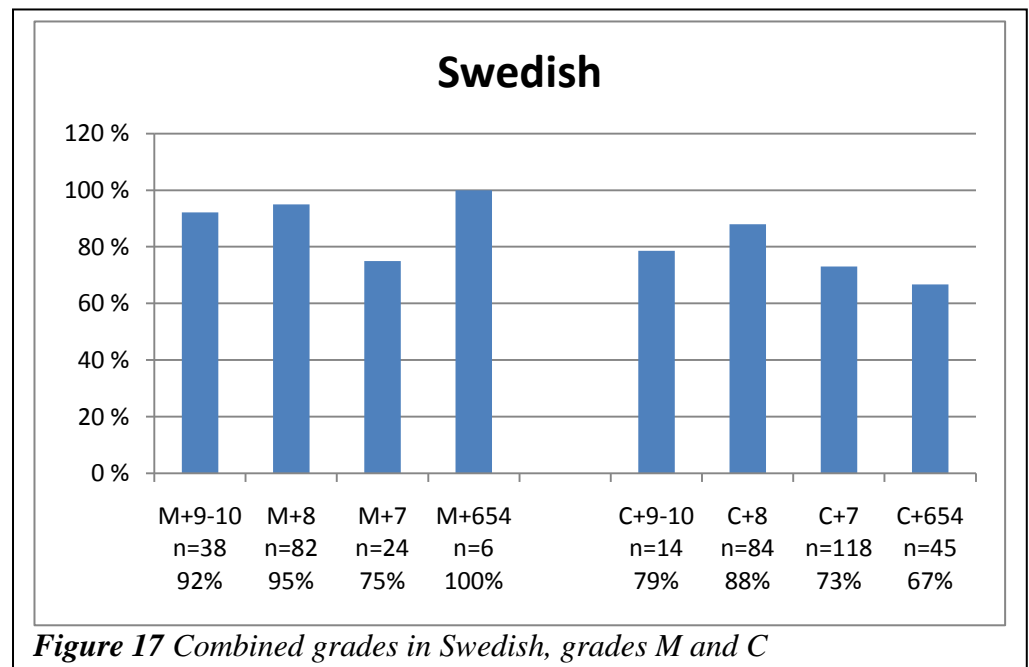
When looking at the figures 13 and 14, and comparing them to figures 2 and 4 from Metropolia it becomes obvious that the results are very much in line, and the differences are only some per cents.



In combined grades the trend is very similar to the results from Metropolia. The highest grades combined with higher final grades pass with utmost confidence, but when the final grade is in the lower end of the scale the level of students passing is much lower. Unfortunately the data sample is that much lower from the other UAS's that some of the combinations become quite rare and thus the validity suffers from small samples, i.e. grade M with final grade 6 or lower has a sample of only seven students, out of which 57% pass, meaning four people. A single student more passing would have risen the level to over 70%, so the significance of a single entity is too great.



In lower grades the small sample becomes more and more obvious, as grade A does not have any students with final grades better than 7. This on the other hand is in line with the assumption that in Metropolia out of those 18 people spread over four years that fall in to this now non-existing category might have entered their results wrong, in purpose or by mistake.



When looking at the Swedish grades in figures 17 and 18 it can be seen that the small sample plays more and more important part, as

the volumes shrink rapidly. In lower grades B and A the amount of students in each category is very small, often less than 20, so the results are hardly valid in this context. But the general results apply, in English the skill level is quite a bit higher than in Swedish, the variation between students within one Matriculation exam grade is quite large, and lower Matriculation exam grades more often than not also mean lower final grades, so to say that the general skill level determines the both grades and they go together quite often.

## 6 DISCUSSION AND CONCLUSIONS

The purpose of this study was to do a comprehensive study on the results of the diagnostic test held to students entering in Metropolia UAS. Also a point of interest was to determine if the results from previous tests could be used as a reference to decide the skill level of entering students. Such tests include the matriculation examination and the final grade from high schools. Also for better understanding of possible national variations or anomalies in Metropolia's results, a cross-sectional analysis was conducted which included four other UAS's results in comparison against Metropolia's results. Students entering Metropolia have to take the test in two languages, Swedish and English, so results from both languages were analyzed.

### 6.1 General

The study was conducted as a quantitative statistical analysis, and a data was collected from Metropolia between the time frame of August 2005 and May 2009. For this study the field of data had to be narrowed down, so a selection of two tests from both languages and from three different fields was selected, giving a total of six tests per language, and a total sample of 6783 for Swedish and 5573 for English. The fields selected were *Technology*, *Health care and Social Services* and *Culture*. The fourth field, *Economics and business administration* was left out because of the lack of results from previous years. From the other four UAS's similar data was asked and collected total sample of 3756 for the school year of August 2008 – May 2009.

At the beginning of the study a problem occurred with the background information that forced a decision to focus mainly on the High School students because their background information was solid and in line, whereas the background information from the students coming from vocational school background was untrustworthy and incoherent. The big picture between the two possible background options was able to draw, and in general the students entering from vocational school

background do not possess similar skill levels as those entering with high school background. Further analysis of the vocational school results was nigh impossible, and thus the focus was mainly on the high school students after this. Also, whilst comparing the results of both matriculation examination grades and final grades, the different nature of these grades needs to be understood. The final grade of a student builds up over several years of high school, and is a compilation of several grades of several courses taken over the years. Students have the ability to choose over additional courses which also makes the structure of the final grade vary between students. These courses are all graded by the accountable teacher alone, and in larger high schools usually student attends courses from different teachers over the years. The teacher, whilst instructed to remain objective over students, might be affected by the student's behavior in class or perhaps the student's results from previous courses. On the other hand the matriculation exams are graded centralized, and anonymously over the Matriculation Examination Board. This ensures that the student's exam is graded solely by the said test, and nothing else. Whilst the final grade builds up over the years, and is not so vulnerable against a singular bad day, the matriculation exam is more sensitive for momentarily weaker performances. Which of these two then represents the students' skills better, cannot be determined, so they are both included in the study, alongside each other and solely.

## 6.2 Results

There are differences in both languages, but especially in Swedish the lower level of skills is apparent, with the worst example being from technology where only 11% of the students in year 05-06 passed the test if entering from vocational school, compared to high school background where the passing level was 39% in said year. The overall level of passing grades was higher in English, even though the demanded level is higher in English, B2 compared to B1 in Swedish. This is a nationwide trend throughout UAS's. Altogether in both lan-

guages over 50% of students do possess the required level when entering UAS, but only slightly over. The problem is the vast differences between fields of study, which is in touch with the student material entering different fields; a fact is that students starting in Culture department have had better grades in High School, and therefore are better equipped before entering the UAS compared to students in the field of Technology. This is of course also due to the fact that the amount of students accepted in Culture department is roughly one fifth of those in Technology, so the entry process is also stricter.

When determining the correlation with the previous grades and the current skill level a three dimensional comparison was made. Firstly the results from matriculation examination were compared against the results from the diagnostic test, then the final grades and finally the combination of matriculation exam and final grade was compared against the result from the Metropolia test. As suspected beforehand, the students with grades on either extremes of the scale did as they were supposed, students with either L or E from matriculation exam, or 10 or 9 as final grade passed the test with almost 100% certainty, as the students with either A or B from matriculation exam, and 4 to 5 as final grade failed with almost as certainly. The most interesting body of students was the one in the middle ground that posed the greatest interest to see whether or not there could be made some major differences within these grades. To see this a combination of these previous results was studied, and the result was that a single grade was not enough to determine the accurate skill level, as especially within the grade C in matriculation exam the divide was significant, as students with C and a high final grade passed almost certainly, the students with C and a lower grade, such as 6 or 5, passed only half of the time. Therefore the mass of students with C cannot be processed as one unified field of students. This tells that if considered the role of this diagnostic test in the future, students' previous grades should not be viewed as singular grades, but as a combination of all possible previous grades, such as the final grade and the matriculation exam

grade. Also if these tests were to be held also in the coming years, an option where the students with the highest grades from high school were to be excluded from the test needs to be considered. Regarding the research question whether or not the students possess the required level of language skills, the answer is never so easy to say. Basically a certain group of students seems to be lacking the required skill level when entering the UAS, and considering the amount of education that the UAS's are able to provide within the four years it would be easy to jump to the conclusion that perhaps some students graduating are falling behind the national requirements. Of course there are numerous factors that affect the differences between the student body entering the school and graduating from it. For one, the UAS's in Finland have a relatively high drop-out percentage, almost one in ten students entering never graduate (stat.fi) and the differences between branches within one UAS, and some engineering branches have a drop-out percentage close to 30%. This quite large amount of students obviously affects the skill level of the graduating students. This would be a perfect subject for additional study; to see what group of students the UAS's are losing during the four years, is there a certain group of students, or are the leaving students spread evenly between the background options.

### **6.3 Recommendations, suggestions & personal thoughts**

When looking at the results of the diagnostic tests and comparing them against the grades of the students, few points stand out relatively clearly. Firstly the necessity of the testing of all students is to be considered. In the light of the results of the diagnostic tests, there are a vast number of students that already when entering the UAS possess a skill level far beyond the required, and this also is apparent from the previous grades these students have received; a high school student graduating with the highest matriculation examination grade inevitably has a deeper understanding in any foreign language. This applies perhaps even more directly to those students whose final

grade from the high school is of the highest level, as that usually means that their knowledge has been good throughout the three years in high school building up that final grade. Thus it is easy to suggest that the students who enter the UAS with a matriculation exam grade of L or E, two of the highest grades, or have a final grade from high school either 10 or 9, again the two highest grades, to be excluded directly from the testing. This sort of upfront elimination should also be done on the other end of the scale, to the students with grades on the lower end of the scale, as they seem to fail the test with similar consistency as the higher grades pass. Therefore it is easy to suggest that the students with the lowest matriculation exam grades and final grades are directed directly to some additional preparatory courses if possible to enable them to achieve the desired skill level for the time of graduation. It is inane to subject these students for testing as the result of the test is blatant upfront.

The most productive student mass to test is the hazy middle ground of the students, as the variation within these grades is quite reasonable, and it would be beneficiary to sort out the students that are sufficiently skilled compared to those in need of additional preparatory coursing. What then should be the limits, between these three groups, ones excluded from testing, those tested and last group of students excluded from testing and directed to preparatory courses? The most troubling group of students are the ones with a matriculation exam grade either C or B as students with grades higher than C pass the test with over 90% certainty, and of students with a grade lower than B less than one in five pass, thus directing them to preparatory coursing would be beneficiary. In final grades this troublesome group is the students with a grade 7 or 6, as the students with grades higher again pass with over 90% certainty, and students with lower grade than 6 fail with almost similar certainty. Oddly these limits apply to both Swedish and English, although beforehand it would have seemed that results in English would have been better. This is partly due to the fact that in English the general level is higher, and the distribution between

grades is different, but the level of actual knowledge correlates well with the grade. Thus there are more students with higher grades in English than in Swedish, but the level of skill with similar grades is equal.

The troublesome group for this study has been since the beginning the students entering with a vocational school background. Mainly the grades they receive posed a practical issue regarding the validity of the results, and had to be discarded because of that, but also when skimmed the rudimentary results that never made into this thesis, the results were baffling. The students are perhaps even more evenly distributed across the board than the students from high schools, and contrary to the presumption, some of the students had notable skills in both languages. The level of language education in different vocational schools varies greatly, and therefore produces also students with very variable skills entering the UAS's. One useful improvement would thus be to standardize both the education but also the grading of such subjects that have effect on the following education, such as the language education. This applies not only to students from vocational schools, but also the small group of students with double degree, vocational school and high school combined. This group of students is troublesome, as they not only give colorful background information when asked, but also in a way distort the results of the vocational school students, as their education is not from vocational school, but results do not differentiate this.

The actual need to test the students can be argued of course until the end of time, but it is a useful tool to determine the need for additional education and also the needlessness of it. It is vital to keep the students from getting frustrated to get their maximum potential out, and both useless revision of already known subjects and also too advanced subjects for one's personal skill level can harm the motivation of the student and hence impair the learning results. Therefore there is a real need for testing of certain groups of students, to determine

the suitable path for the oncoming years in UAS. What needs to be discussed is the amount of language education in UAS's, and its sufficiency. The basic setting of one three credit course of both languages spiked with perhaps one another preparatory course is sometimes not enough to educate students to a sufficient level of language skill. If the goal is to meet the requirements dictated in the decree considering the level of language understanding of students graduating from UAS, there should be more, deeper and diversified language education available for the students. Unfortunately this is a question that goes far deeper and beyond the limits of one thesis project, so therefore it should not be discussed further, but be left as a notion for the reader to process and treated as seen suitable.

## References

- ARENE language group. Language teaching practice recommendations by the ARENE, Rectors' Conference of Finnish Universities of Applied Sciences language group (ARENEN kielityöryhmä. Kieltenopetuksen käytäntösuositukset) Dec 12, 2006. Available at: <http://extra.seamk.fi/arenekr/kaytantosuositukset.htm> [Accessed Dec 18, 2008]
- CEFR - *Common European Framework of Reference for Languages: learning, teaching, assessment* (2001) Cambridge: Cambridge University Press. Finnish translation: Eurooppalainen viitekehys: Kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys (2003) Council of Europe. Helsinki: WSOY.
- Universities of Applied Sciences Act/ Ammattikorkeakoululaki . Available in Finnish at: <http://www.fiblex.fi/fi/laki/ajantasa/2003/20030351> . [Accessed Jan 8, 2009]
- Universities of Applied Sciences Decree/ Ammattikorkeakouluasetus. 352/2003 §7-8
- Kantelinen, R. & Airola, A. (2007) Ne antavat yhtenäisempää taustatukea kieltenopetukseen Kokemuksia ammattikorkeakoulujen kieltenopetuksen käytäntösuosituksista (They provide more uniform support for language teaching - Experiences in ARENE recommendations of language teaching) Joensuun yliopiston selosteita. Joensuu: Joensuun yliopisto. Accessible at: [http://joypub.joensuu.fi\\_publications\\_frameappl\\_pdf.pdf](http://joypub.joensuu.fi_publications_frameappl_pdf.pdf) [Accessed Jan 10, 2009]
- Jaatinen, P (2005) Ammattikorkeakouluopiskelijan kielitaito – Harhaa vai Hallintaa? (Kieltenopetuksen kehittämisprojekti – Seurantatutkimuksen 3. osaraportti / Development project for language education, follow-up study report number 3.)