

IMPROVING THE COMPARABILITY OF GRADES

Background

Professor Arvo Lehtovaara was a member of the Matriculation Examination Board and acted as its chairman from 1967 to 1972. He saw that the level of difficulty of the matriculation examination varied from one year to another, causing problems in the criterion-based evaluation. This fluctuation was of relatively minor significance as long as the completion of the matriculation examination guaranteed a study place at a university but, as an ever-growing number of students started to look for a study place in the years to follow, it was no longer possible to compare different results.

The development of information technology enabled the transition to the two-scale evaluation system still in use today. Each examination was first evaluated on the basis of several criteria. After this, the student-specific total score was placed inside a relative grade framework based on a normal distribution. The primary objective of the relative evaluation system was to have properly comparable examinations, but the system also offered a degree of comparability between different subjects because students had only little impact on choosing the tests included in the examination.

In the current matriculation examination system, only the mother tongue test is compulsory for all. Increasing the level of options in the examination means that there are tests taken by a highly selected group of students. If a normal distribution is used as the basis for evaluating for these tests, the students will obtain too low a grade considering their skills. As a result, grades in different subjects, or even in different tests of the same subject, are not comparable. However, comparability of grades is a prerequisite for being able to use matriculation examination grades reliably and fairly in selecting students for universities and universities of applied sciences.

Average of standard total scores

The Matriculation Examination Board has decided to improve the comparability of grades by using the average of standardised total scores (SYK) in defining the score limits.

Two successive examinations are used in the selection of the population, i.e. the students of one full year.

Standardisation is required before SYK can be used.

For this purpose, the Z-value of each student is first calculated in each subject.

$Z\text{-value} = (\text{score} - \text{student average}) / \text{dispersion}$,

Score = a student's examination score

Student average = the average score of all students from a test in a single subject

Dispersion = dispersion within the specific test

The Z-value is calculated for each student for each test the student has taken and during each examination period. Using the Z-value, tests of different subjects are comparable.

SYK is calculated as follows:

$$\text{SYK} = (zA + zB + zC + \dots) / \text{the number of subjects included in the examination}$$

z = Z-value

A, B, C = tests of the subjects included in the student's examination

When calculating the SYK, any tests taken by the student during the examination period under review or during previous examination periods are also taken into account. If the student later adds new tests to the examination, the SYK figure will change.

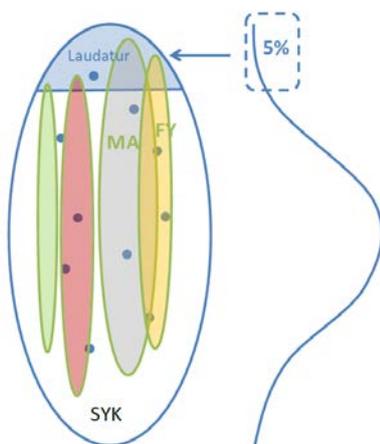
Selection of population

If the population is selected according to examination period, there will be problems because the groups of students attending the examination in spring and autumn are highly different. As a result, the population is selected using two successive examination periods. The spring examination includes the previous autumn and, correspondingly, the autumn examination includes the previous spring. In this way, the SYK figure of students from a full year can be calculated.

Using the SYK figure

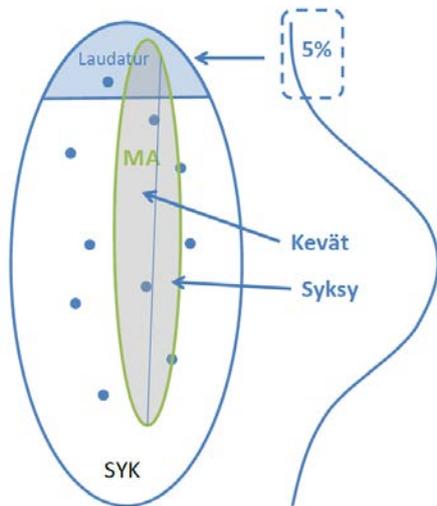
Using SYK figures, the score distribution can be assumed to follow the normal distribution within the full-year population. Therefore, it is possible to grant laudatur grades to 5% of all full-year students.

Using SYK, the grades of different subjects are comparable by reviewing the SYK figures of all students taking different tests during the year.



It is also possible to review students taking different tests among all full-year students. In this case, the portion of a designated grade of a specific test (e.g. laudatur) is based on the SYK figures of the specific

group of students. Furthermore, the grades of a specific test are comparable between autumn and spring examinations. In the image below, the advanced mathematics test is selected.



The SYK figure will be adopted in stages. In the spring 2014 matriculation examination, the SYK figure was used to define score limits for the tests of advanced mathematics, physics, chemistry, advanced Finnish, advanced German, advanced French, advanced Spanish, basic German and basic French.

Improving the comparability of matriculation examination grades is aimed at better utilising the matriculation examination when selecting students for universities.