Thematic approaches to equality and equity in basic education

Jorma Kuusela
Finnish National Board of Education
Assessment of learning results, 6/2006

SUMMARY

This report combines ten assessments of learning outcomes in the final 9th grade of basic education carried out by the Finnish National Board of Education (FNBE) between 1998 and 2004. These include four mathematics assessments, three mother tongue and literature assessments and individual assessments of natural sciences, A-syllabus English (foreign language starting at primary level) and B-syllabus Swedish (foreign language starting at lower secondary level). In addition to sample schools, the analysis also includes the learning outcomes of pupils at those schools that had ordered[1] assessments since 2000. Consequently, the data covers more than 90,000 pupils, equating to 1.5 times the size of the entire age group.

The purpose of the research was to:

1) shed light on regional variations in average learning outcomes and the statistical correlation between average learning outcomes and the socioeconomic variables describing the backgrounds of pupils at different schools;
2) study the equitability of pupil assessment; and
3) reflect on issues relating to provision and allocation of support measures.

In other words, all themes are related to equality and equity in education in one way or another.

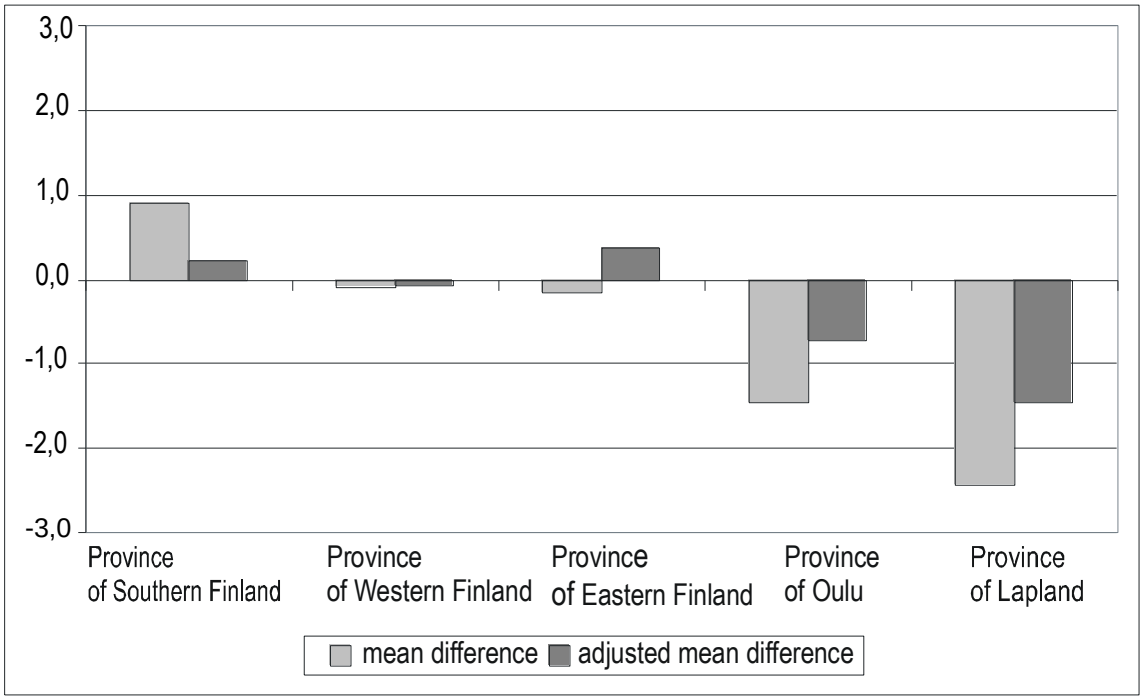
Variations in learning outcomes

Regional variations in learning outcomes are examined according to major regions, such as provinces, EU development areas and the statistical municipal classification. The general finding is that there are no considerable differences between major regions when measured in terms of pupils' average performance, and that the situation becomes even more balanced when socioeconomic background factors describing schools' operating environments are taken into account. The unit used in the following figure is comparable with the rate of correct answers in national assessment tests. The national average is set at zero.

The remaining differences between major regions may be attributed to sociological rather than pedagogical factors. Attention has focused on aspects such as the weak scores among boys living in the Province of Lapland in assessments of the subject known as mother tongue and literature. A more in-depth analysis shows that this only concerns pupils in the province's rural schools. Similar results have also been found in Northern Sweden and Iceland, albeit in assessments of mathematics skills. Nevertheless, there is no reason to assume that this phenomenon – dubbed *the Jokkmokk effect* after a

[1] Education providers can commission the FNBE to carry out evaluations when they are not included into national assessments. The education providers pay for this service,

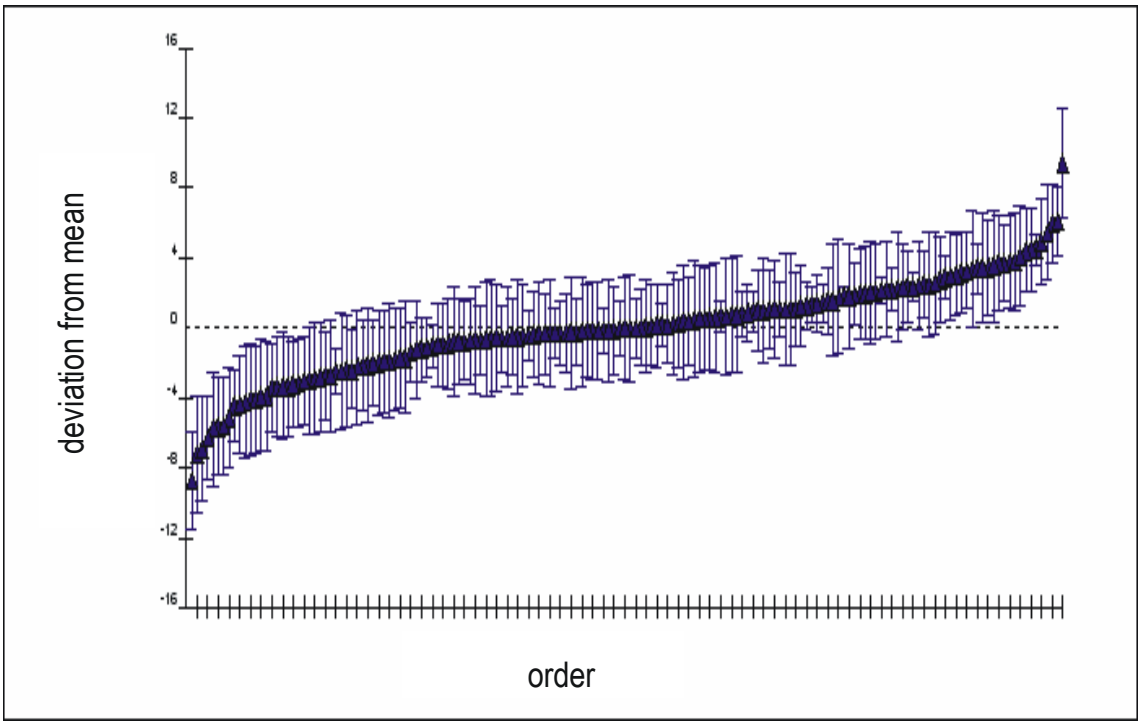
remote rural area of Northern Sweden – would only apply to mathematics; this is more about the fact that the phenomenon has been easiest to detect in mathematics assessments. It is noticeable when girls achieve significantly better results in mathematics (or physics) than boys, whereas this is expected in other subjects.



Summary figure 1. Differences between provinces in average learning outcomes in general terms and when adjusted by socioeconomic background data (averages 50, deviations 20).

However, different methods may yield different results; by way of example, an earlier analysis conducted by Jakku-Sihvonen and Komulainen in 2004, which was based on ranking schools into quartiles, leads to a different interpretation about the differences between major regions and, in particular, between Northern Finland and the Helsinki Metropolitan Area, than comparisons based on averages. With the latter, even regions are too large as units to allow detection of any significant differences in learning outcomes. Moreover, Jakku-Sihvonen and Komulainen did not have access to socioeconomic data concerning pupils from different schools. When factors describing the average socioeconomic backgrounds of pupils at different schools and, in particular, those concerning parents' educational backgrounds are taken into account, the results of the Province of Lapland, for example, are in line with a prognosis that can be made on the basis of pupils' average socioeconomic background, even when analysed in terms of quartiles. In other words, it is difficult to find any systematic differences at the level of major regions that would indicate deficiencies in education provided by schools or within the education system.

In this report, differences in pupils' average test scores become evident at municipal and school levels. Differences between the extremes are considerable, but the majority of municipalities and schools are placed within the 95percent confidence interval from the national average.



Summary figure 2. Averages of pupils' scores and 95% confidence intervals in those municipalities (n=164) with at least 100 pupils involved in the assessments.

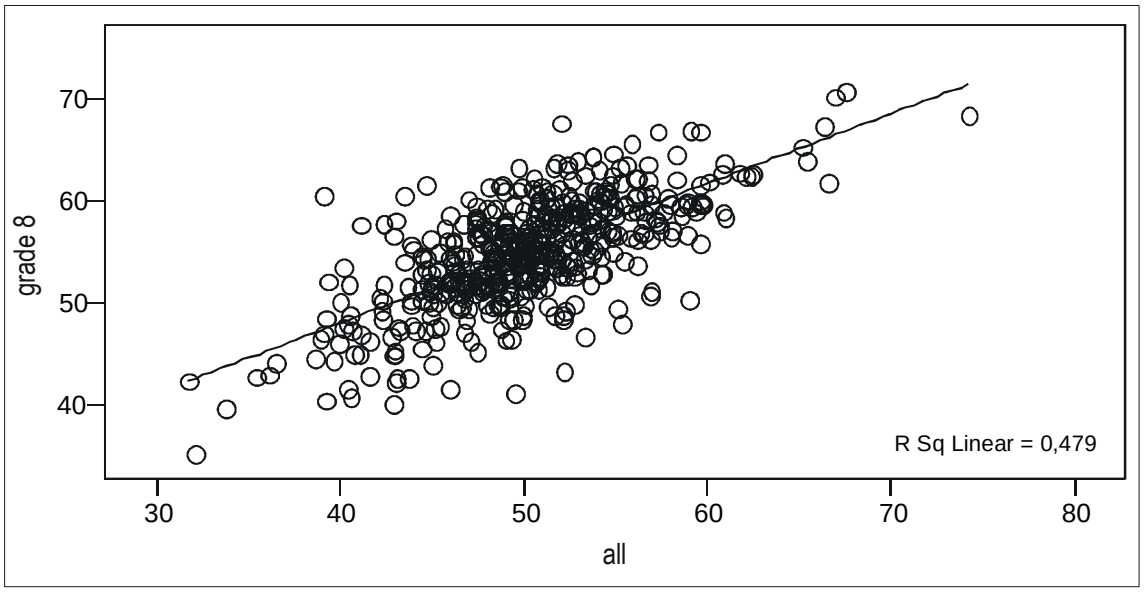
There is a statistical correlation between the average socioeconomic backgrounds of pupils at different schools and schools' average assessment results. Thus good learning outcomes are related to a high average level of education and a low rate of unemployment. However, all indicators of social advantage correlate strongly with each other both at school and municipal levels. This link is even more pronounced at school level, as inter-school differences are highest in large urban municipalities. This means that the results of major cities tend to set the averages in a municipal analysis.

Socioeconomic background variables facilitate a prognosis for schools' average performance in assessment tests. The prognosis works better with urban schools than rural schools. This can partially be attributed to statistics, because urban schools are generally larger and have requested assessments more often than their rural counterparts. This leads to a smaller random variation, which means that the role of pupils' background factors becomes more pronounced. Another technical reason is the fact that differences between schools are larger in urban than rural areas in terms of both socioeconomic background variables and average learning outcomes. There are no material differences between average learning outcomes of pupils at urban schools and rural schools. It appears, however, that individual variation in learning outcomes, which occurs within schools in rural areas, becomes more of a case of inter-school variation in major cities. In addition, pupil selection increases inter-school differences, regardless of whether this is examined in terms of pupils' average test performance or indicators of pupils' average socioeconomic backgrounds at different schools. Nevertheless, there are also variations in learning outcomes between rural schools that can neither be explained by means of the background variables used in this context nor by any technical reasons.

Equitability of pupil assessment

Pupils' grades vary by subject in a manner that cannot be considered to conform to the assessment criteria of the National Core Curriculum for Basic Education in all respects. The difference in grade distribution is highest between B-syllabus Swedish and mother tongue and literature. Pupil assessment complies with the policies set out within the National Core Curriculum in that those who have performed better in their assessment tests have usually received better school grades as well. While there are some individual schools where grades are only weakly correlated to national assessment test scores, intra-school comparisons also indicate that pupil assessment is mostly in line with the criteria of the National Core Curriculum.

Conversely, inter-school comparisons yield different results. The grade variation between schools may be as high as 1.5–2 grades among pupils with equal average levels of performance in assessment tests. The most important factor explaining this variation is the schools' overall performance level. Grades are adjusted to the school's own overall level rather than the criteria issued by the Finnish National Board of Education for competence for grade "good" (8), for example. In general terms, the average rate of correct test answers among pupils awarded grade 8 (good) on their school reports varies by school, such that it is slightly above the school's average rate of correct answers – irrespective of whether it was 40percent or 80 percent. The fact that pupils were asked about their most recent rather than their final report grades in connection with assessments does not explain such a systematic variation, as the grade distributions obtained from pupil surveys are very much in line with the distributions of final grades extracted from FNBE register data.

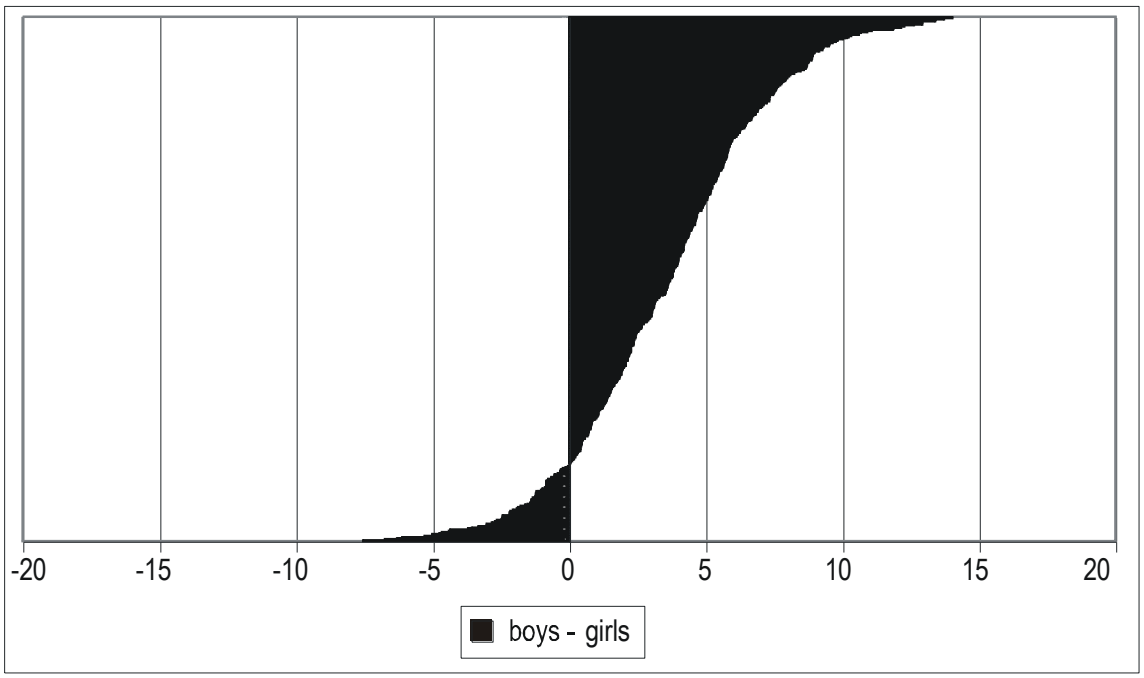


Summary figure 3. Correlation between averages of standardised rates of correct answers among all pupils and those awarded grade 8 in schools with more than 10 pupils with grade 8 (all assessments).

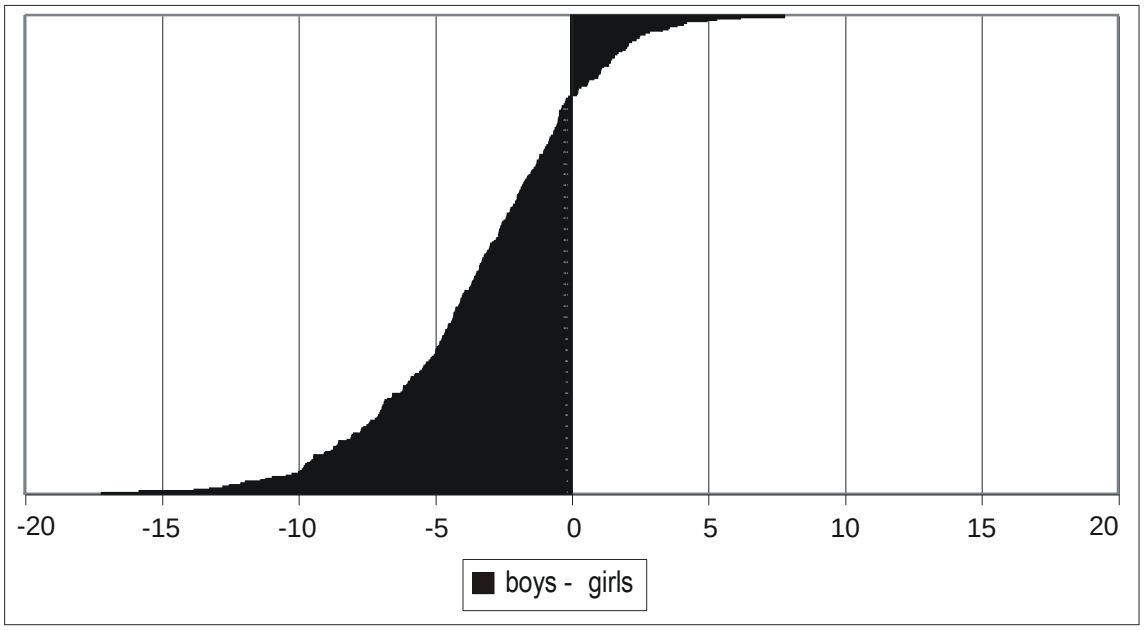
The figure above only examines the average performance of pupils with grade 8, with assessment tests standardised for consistency. In order to reduce random variation, the analysis only includes schools with more than ten pupils who have participated in an assessment and received grade 8 on their reports. The key message to be taken from the figure is that about half of the grade variation can be attributed to the school's overall performance level. Grade 8 was used as the point of reference, because the Finnish National Board of Education has issued criteria for competence for grade "good", which means grade 8. By using a slightly more complicated method allowing analysis of all grades, the correlation between a school's overall performance level and its assessment practices increases even further. The same also happens when examining one subject at a time.

It is also possible to detect other systematic differences in pupils' grades. Assessment practices applied to boys and girls comply with gender stereotypes: on average, boys are required to perform better to receive a certain mathematics grade than girls, while the situation is the opposite in mother tongue and literature. While this could be attributed to reluctance to systematically award weaker grades for boys in mother tongue and literature, the explanation does not apply to mathematics, which involves virtually no gender differences in learning outcomes. Neither do pupils' attitudes towards the subject provide an explanation, because differences in attitudes should have the opposite effect. These differences are not considerable in average terms, but they are systematic and, in some schools, so high as to raise questions about equality of treatment of pupils.

Figures 4a and 4b compare boys' and girls' grades in mathematics and mother tongue and literature by school, taking performance in FNBE assessment tests into account. A column pointing to the right indicates that girls have received better grades than their assessment test scores would suggest, while a column pointing to the left means that the same is true for boys. The unit used here is the standard score that can be equated to percentages.



Summary figure 4a. Gender differences in all mathematics assessments by school and in those schools with at least ten boys and ten girls involved in an assessment.



Summary figure 4b. Gender differences in all mother tongue and literature assessments by school and in those schools with at least ten boys and ten girls involved in an assessment.

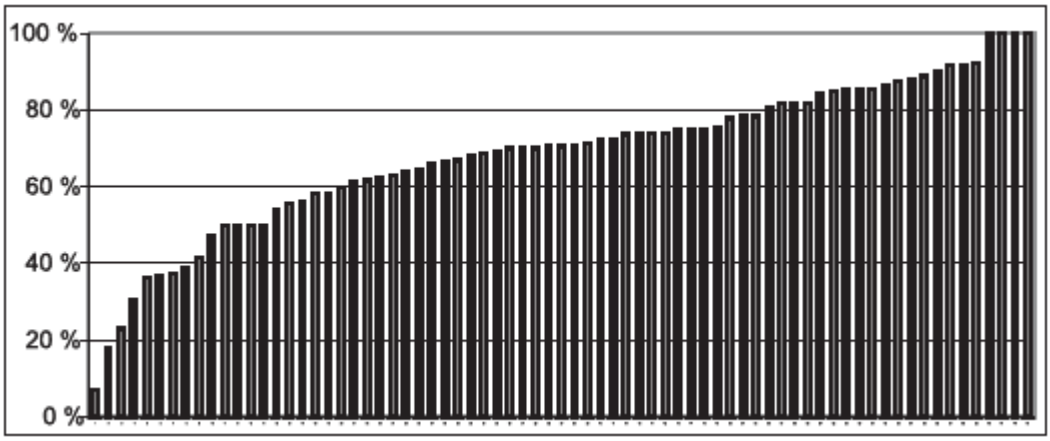
Another difference is related to pupils' future plans. Pupils planning to go to upper secondary school are assessed on more stringent criteria than those choosing another option, which mainly means vocational education and training (VET) and, in some cases, the voluntary additional 10th grade of basic education or a gap year. Again, differences are not very significant in average terms, but they recur in all assessments included in this analysis and they can also be considerable in some schools. This cannot be explained by the possibility that pupils in high-performing schools would be more likely to continue to upper secondary school, while a high performance level would lead to stricter assessment. Neither can it be attributed to the fact that a higher proportion of girls than boys move on to upper secondary school. In other words, the explanation needs to be identified using non-statistical methods – namely, through a qualitative approach.

Equitability of support provision

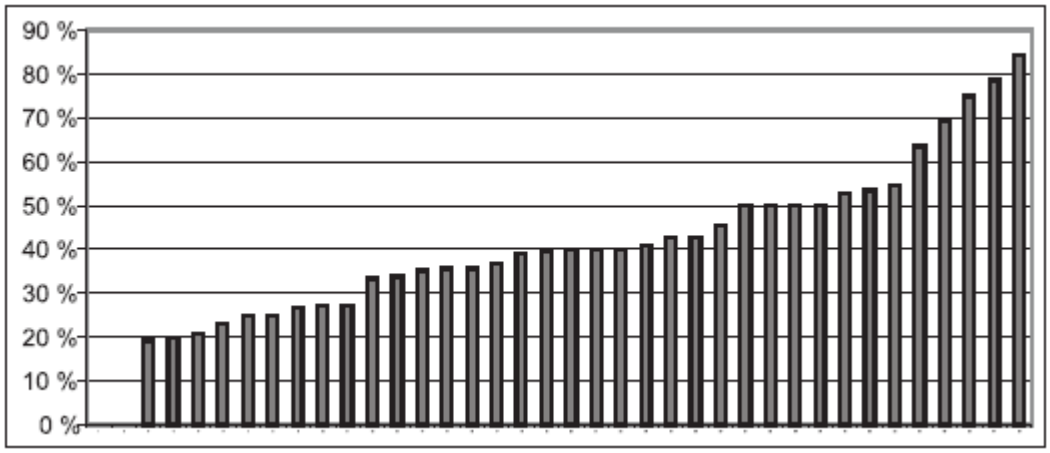
In terms of support measures mentioned in the Basic Education Act, this analysis has focused on those assessments that also asked whether pupils had received remedial instruction or special needs education. Assessments have been conducted on pupils studying according to the mainstream curriculum, which means that the research population does not cover pupils admitted to adjusted special needs education, for example. Those special schools providing education according to the mainstream curriculum have also been excluded from the data. Information obtained from pupils differs from figures provided by Statistics Finland, for example, in two ways. On the one hand, it can be considered as being the kind of retrospective data that describes pupils' history, rather than being cross-sectional data from a specific statistics collection period. On the other hand, this data is capable of linking pupils' performance in assessment tests to whether or not they received support. Comparisons are complicated by the fact that the wording of questions and possible responses has varied and by different limitations used in different assessments. The reason for this is that support measures have been included as background variables rather than being the actual focus of analysis in assessments, whereas the roles of variables have been reversed here. Nevertheless, it is possible to draw some general conclusions.

In terms of all pupils, less successful ones have generally received more support than those who have done well, which is quite natural. There are neither considerable nor systematic differences between major regions, such as provinces, or different types of municipalities (city, town or rural municipality). However, there are very significant variations in provision of support between municipalities. These differences remain considerable and even increase in some cases, even if the situation were only examined in terms of pupils falling within the lowest-performing quintile. While the 20 percent cut-off point is a generalisation, pupils in this group have achieved relatively poor scores in all assessments. In cases where support has been provided, it has mainly been allocated consistently within municipalities in the sense that pupils in the lowest quintile have received more support than average.

As an example of municipal variations, it is possible to describe support received by the lowest-performing quintile in mathematics and Swedish (B syllabus), as these represent the best and weakest subjects in terms of support provision. The assessment of learning outcomes in mathematics included a question concerning whether pupils had received remedial instruction during the entire period of basic education and part-time special needs education during the final 9th grade, whereas the assessment of learning outcomes in Swedish only asked whether pupils had received remedial instruction. The different formulation of questions does not distort the results to any great extent due to the small proportion of pupils reporting to have only received special needs education in mathematics. The precedence of remedial instruction may also be considered to represent the natural order of support measures, with the exception of speech therapy and dyslexia support for mother tongue and literature.

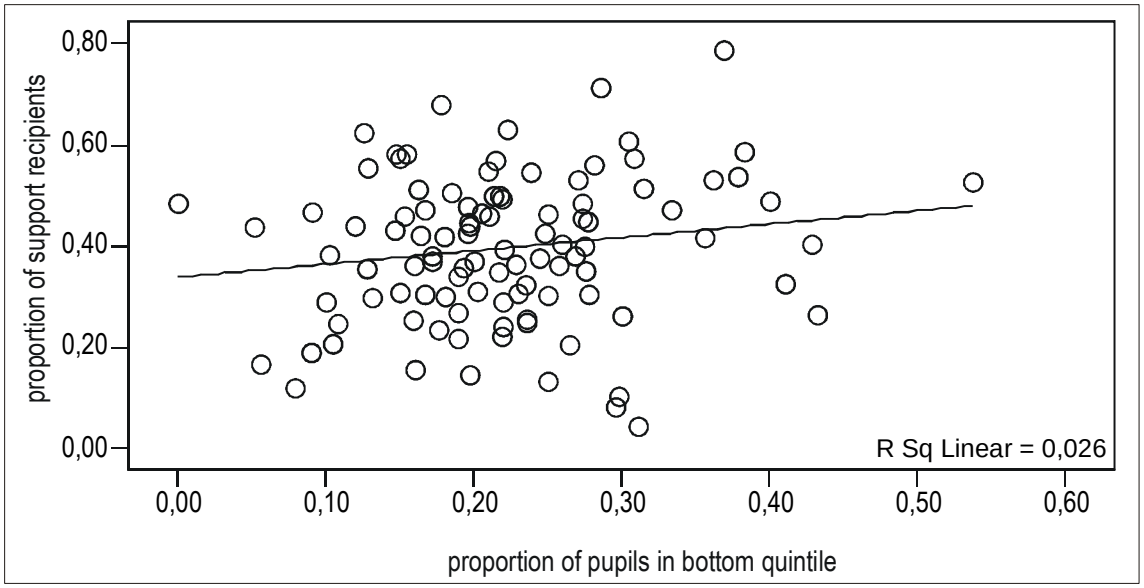


Summary figure 5a. Proportion of pupils in the lowest quintile who had received remedial instruction and special needs education in municipalities with at least 10 pupils falling within the lowest quintile. Mathematics assessment, 2002.



Summary figure 5b. Proportion of pupils in the lowest quintile who had received remedial instruction in municipalities with at least 10 pupils falling within the lowest quintile. B-syllabus Swedish assessment, 2001.

As the lowest-performing pupils have generally received more support than their peers both when examining the entire national data as a single group and when conducting an equivalent analysis by municipality, this is the expected correlation trend. It would therefore seem logical that provision of support would generally be more prevalent in those municipalities and schools with the weakest average learning outcomes. If this were the case, there would also be a match between demand and supply when comparing municipalities and schools. It would also be possible to find a natural explanation for an opposite link: timely support has helped pupils overcome their learning difficulties. In reality, however, the situation is quite haphazard. While pointing in the right direction, the correlation coefficients are low and the proportion of variance explained by the coefficients varies from just below 1 per mille up to 14 percent. The results are not affected by whether demand is assessed in terms of pupils' mean performance or based on the proportion of those in the bottom quintile. In addition, the results do not change even when attempting to allow for the amount of support received by pupils; nor do they change materially even when examining schools rather than municipalities; and they do not change even when examining remedial instruction and part-time special needs education separately. While variation in questions was mentioned as being a disadvantage, it is also possible to turn this around and add that the situation remains the same regardless of how the question was framed. When assessed by comparing municipalities or schools, the link between demand and supply mainly appears to be random. The reference data used here is the 2002 assessment of mathematics learning outcomes. There is a correlation between the proportions of pupils in the lowest quintile and those who had received support, which explains less than 3 percent of variation in provision of support.



Summary figure 6. Proportion of pupils in the lowest quintile and proportion of pupils who had received support in municipalities with at least 30 pupils involved in the assessment. Assessment of mathematics learning outcomes, 2002.

It is also possible to detect a certain gender effect in mathematics in that, considering their performance levels, girls receive more support than boys. This may be partially attributed to the fact that boys are less willing to participate in remedial instruction than girls, at least when it comes to mathematics. This, in turn, may be linked to the fact that boys overestimate and girls underestimate their own abilities in the subject. While equally pronounced differences in the opposite direction cannot be found in mother tongue and literature or in Swedish, it is not completely out of the realm of possible explanations that the stereotypical idea of boys having a 'better knack for maths' – which may be an underlying reason for the differences in pupil assessment – may also affect provision of support.

The statutory right of pupils to receive remedial instruction or special needs education appears to involve problems. The most significant problem are the differences between municipalities in provision of support that cannot be explained at least by differences in performance as per FNBE assessments. A good point here is that we cannot go as far as proving that a lack or inadequate provision of support may actually be compensated by relaxing pupil assessment.

No national statistics whatsoever are collected for remedial instruction. Provision of special needs education, in turn, is described by means of numbers of pupils in special needs education, rather than in terms such as lesson quotas relative to the numbers of pupils in part-time special needs education. The main problem of research and public debate appears to be related to the fact that the real demand for support is not known and compilation of related statistics seems to be impossible. When people point out the increasing demand for special needs education in public debates, they actually describe the increase in terms of provision. Even if provision were always based on local assessment of demand, it appears that these assessments vary considerably in different municipalities.

Conclusions

The common conclusion derived from these analyses is that, even though the quality of the Finnish education system can be considered as being exceptionally consistent in international terms, the education policy decisions of municipal authorities and the practices of individual schools play a significant role. There are variations between schools and municipalities in terms of learning outcomes, pupil assessment and support received by pupils that need to be reduced. Responsibility for supporting education providers and schools lies with the Finnish National Board of Education and the key question is how it could better fulfil this role without compromising the decision-making powers of municipal authorities and schools.

The finding that the analysis of average scores does not reveal any considerable differences between *major regions*, even without correcting them by variables of pupils' average backgrounds, can be considered as being a positive result. In addition, these differences are further reduced when allowing for socioeconomic factors, with their unequal regional distribution. However, different methods lead to different results and interpretations and this definitely does not make the work of education policy-makers any easier. Even so, information about the effects of methodology and intervening variables will probably provide a more diverse overview in support of the decision-making process.

Introducing another explanatory variable does not, however, solve the actual problem. It is not easy to break the chain of cultural heritability of education (in the statistical sense of the term), and no country has been completely successful in this effort. The Ministry of Education's Development Plan for Education and Research for 2003–2008 cited a special risk of so-called inter-generational exclusion, in which exclusion is passed on to the next generation as a result of social and cultural deprivation. In this respect, attention needs to turn to municipalities and schools, which reveal differences that may be seen to be considerable in some extreme cases. The challenge involved in preventing this risk is demanding, but it must be accepted and Finland's relative success in achieving equal opportunities in education puts it in a better position to do so than virtually any other country.

Increases in resources are offered as being a universal solution to almost all problems, but one may still – without questioning the demand for extra resource allocation – challenge a summary and untargeted increase, because this is a very expensive solution. As schools and municipalities have very different starting points, the kind of positive discrimination that is already being used in Helsinki, for example, is a possibility for major urban municipalities. Supporting small municipalities through targeted extra resources would, at least in theory, be one possible way of aiming to level out average differences between schools in terms of pupils' performance. Another – and probably a more efficient – possibility would be systematic development of those support systems that can improve the learning abilities of pupils with the most disadvantaged circumstances. National programmes promoting learning to learn and other general skills might also function as a tool for differentiation, while contributing to achievement of those objectives relating to equality and promotion of intellectual growth and learning that have been mentioned in schemes such as the Ministry of Education's Strategy 2015 programme.

It is hardly possible to overemphasise the significance of early identification of support needs; one possibility would be production and distribution of age-appropriate diagnostic materials intended for everyone on the FNBE website, for example. The Ministry of Education and the Finnish National Board of Education are currently developing the schools' quality systems, which may contribute to encouraging earlier and more systematic diagnosis of learning difficulties, proper quantification of support needs and consistent provision of timely and adequate support measures. There are many possibilities available, but deciding on their selection and implementation is an education policy issue.

The results and subsequent conclusions presented in this report have not been unforeseeable. People have known to look out for these and have already reacted to areas in need of improvement. Increasing attention has been focused on pupil assessment by determining criteria for good competence for earlier transition points, in addition to final assessment. This, in turn, has made it possible to anchor assessments of learning outcomes to the objectives set out within the National Core Curriculum. Continuous improvement of national assessments and feedback intended for schools and education providers, complete with increasing their effectiveness, already forms part of the work carried out by the Finnish National Board of Education. The new National Core Curriculum also provides the opportunity to distribute the FNBE assessment tests over transition points, instead of focusing primarily on learning outcomes in the final grade of basic education, and to develop a 'soft' feedback-based steering system in terms of pupil assessment and support measures as well.

Statutory support measures for pupils have also been addressed more clearly within the new National Core Curriculum, while two different development projects, funded by the Ministry of Education and implemented by the Finnish National Board of Education, have been launched, namely, the 'Different Learners' project and the ALPO project geared towards development of subregional co-operation in provision of support. In this respect as well, this report can mainly be considered as being a justification for the focus areas of the National Core Curriculum and as an indication of the fact that the document should also be taken seriously as part of the municipal decision-making processes concerning education.