RITVA JAKKU-SIHVONEN and JORMA KUUSELA

EVALUATION OF THE EQUAL OPPORTUNITIES IN THE FINNISH COMPREHENSIVE SCHOOLS 1998–2001

Evaluation 11/2002
NATIONAL BOARD OF EDUCATION

Layout: Layout Studio Oy/Marke Eteläaho

ISBN 952-13-1610-1 (nid.)
ISBN 952-13-1611-X (pdf)

ISSN 1238-4453

Yliopistopaino Oy, Helsinki 2002

FOREWORD

The basic education[1] reform represents one of the most significant socio-political decisions made in independent Finland.

Probably no other peace-time reform is comparable to it. Of course it is rather arbitrary to compare reform processes, but we can be justified in stating that the basic education resolution is equivalent in significance to the Finnish Primary School Act of 1866. Only a handful of such socio-political turning points can be singled out in Finland's entire civilised history.

One of the central principles of the basic education reform was the idea that a system based on a comprehensive school model could guarantee students' basic civil rights more reliably than the old parallel schooling system. The proponents of the reform wished to ensure that the duration and quality of a student's education would not be conditioned by the kinds of circumstances the student was born into or by the municipality in which his or her studies began. Before the reform, the basic education received by Finnish students varied at random in duration and quality, and access to higher levels of education was a privilege granted more according to the individual's social background than according to his or her talents.

The comprehensive school model was designed to increase the country's pool of skills. Until then, for instance the basics of foreign languages were available only to a part of students. It was possible to produce a shift from a social structure based on primary production, which required low levels of professional skills, to a production structure that was more productive in terms of the country's economy, only by increasing the general level of education and by making specialist knowledge available to a broader section of society.

Our concept of equality has changed during the history of basic education. Towards the initial stages of the reform, the implementation of equality was measured on the basis of the distribution of education places and the extent to which the availability of education was looked after. These criteria were clearly quantitative. Basic education has now fulfilled these criteria. As the reform progressed, the demands for educational equality became stricter. It was not enough that we ensured the equality of students' access to education; equality began to be evaluated in terms of results.

The separation method adopted in the initial stages of basic education, which led to subject levels, resulted in different opportunities for access into further education based on the amount of teaching a student had received in mathematics and a first foreign language between the seventh and ninth grades[2] . The decision conflicted with the new concept of equality and the idea was abandoned at the time of the educational legislation reform in 1985.

1 Basic education in Finland refers to the combination of primary and secondary education, given to students aged between seven and 16.

2 This refers to secondary education in Finland, given to students aged between 13 and 16.

These days educational equality has to be assessed in terms of students' learning-based results. Significant differences in students' levels of performance can adversely affect their choices of subjects for further education and their success in the intermediate grades[3] . Nobody should by now be relying on basic education alone.

It is perfectly natural that differences exist between schools' learning-based results. It is also a characteristic of the education given to students of this age group that there will be differences between individuals' results. We must accept these differences, but it is a task for social politics to decide what size differences we should allow. This relates to the kinds of variations in students' results that can be affected with political decisions. Politics cannot change a student's genetic inheritance, but it can change a student's learning environment. If we can prove the existence in students' results of systematic differences, affected in a significant way by individual schools, the responsibility for rectifying this will lie on the shoulders of our social policy-makers.

The requirements made of our citizens will in future become even more demanding than now. Our knowledge and skills will be challenged more than ever by requirements of a balanced contribution to social life, of developing our skills to interpret reality in an ever more complex world, of securing our livelihood in the working environment and of surviving the routines of everyday life. The wish that new technologies would free us from the pain of learning has not been fulfilled, and we cannot continue to bank on such an unfulfilled dream.

The responsible management of a school implies the nurturing also of students who learn more slowly. It is also justifiable to demand that the differences in results caused by schools should be reduced by trying to increase the level of results in general. Assessments that have been carried out on learning-based results so far indicate that there is a clear connection between a student's results and the operating environment of the school he or she attends. Therefore we cannot eliminate the differences in results with reforms that affect schools alone.

Many exceptional conclusions can be drawn from data related to differences in students' results, and the significance of these conclusions can be evaluated in various ways. The development of schools is never complete: the targets are like a horizon, which appears further away the further you progress. The conclusions must be drawn carefully, with the knowledge that no final solutions are attainable. In my opinion, however, we can, on the basis of the material published here, set a new indicator for the welfare of a district, a municipality or a sector: that of the basic education results of its students.

Jukka Sarjala

3 This refers to education received between secondary and higher education by students aged 16 to 19.

CONTENTS

INTRODUCTION

This publication concentrates on the examination of the ability of Finnish basic education to implement educational equality.

In terms of educational policy, we consider educational equality to refer to the inexistence of systematical differences between girls and boys, between different social groups and between geographical areas. Differences between individuals' results are natural. On the other hand there are various kinds of random or systematical reasons for the appearance of school-specific variations in results, and both didactic and social measures need to be taken in order to rectify this.

School-specific differences can be subdivided into two types, of which one reflects variations in teaching or organisation, and the other, variations in the schools' operating environments and the socially directed result-orientation of the education. Problems related to educational organisation and the quality of teaching can effectively be solved with pedagogic measures. The elimination of differences caused by operating environment and social variables, on the other hand, requires thorough examinations of educational policies, clear declarations of intention and large-scale procedures.

Basic education has been assessed by the Finnish National Board of Education since the beginning of 1995. A broad-ranging evaluation of the state of basic education called "Toteuttaako peruskoulu tasa-arvoa" ("Are Policies of Equality Implemented in Basic Education?") was reported on in 1996. One of its main conclusions was that a certain fall in the standard of results in the students' first language (Finnish or Swedish) could be observed. A planning project related to the system by which learning-based results are assessed was initiated on the basis of these observations on results, of the looseness of the 1994 teaching plan principles and an educational policy that emphasised the operational freedom of schools. Its aim was to develop an assessment system that would provide the authorities and decision-makers with information on the extent to which schools meet the targets set for them in the teaching plan principles. The most important motive for the development of assessment was to support educational equality with a new kind of administrative organisation. The decision-makers and managers have to be provided with reliable data on whether the boys and girls in different social groups all around Finland meet the targets set for basic education.

The first procedures carried out according to this assessment system were initiated in 1998. From the beginning, one of the operating principles for these procedures has been the comparability of the data acquired; i.e. the production of information on trends, which can be used for educational improvements.

The articles in this publication are based on the assessments carried out on basic education between 1998 and 2002. On the basis of these assessments it is

possible to draw conclusions on how well the national aims have been met in various schools that offer basic education.

The aim of these articles is to examine, on the basis of learning results, the extent to which procedures that have been implemented will help schools meet the targets of educational equality. In practice, students' results form the most reliable way of doing this, because the meeting of these targets set the bases for students' access into and success in further education.

The first article examines the geographical and gender-related differences in learning results and attitudes, in both the schools with the weakest performance levels and those with the strongest ones. The second article examines the economic and social characteristics of the municipalities in which these schools are located.

Ritva Jakku-Sihvonen

DIFFERENCES BETWEEN SCHOOL-SPECIFIC RESULTS IN TERMS OF EQUALITY IN EDUCATION

1 INTRODUCTION

1.1 The Purpose of the study

Equality is one of the most important aims of educational politics in Finland, and the concept is based on the premise of each citizen's right to receive a basic education, prescribed in the Finnish Constitution (Section 13). Equality in education is defined as the right of all students to receive the same level of educational services regardless of their gender or place of residence, or the socio-economic status of their families.

Examining the implementation of equality in education by assessing students' learning results is in accordance with the assessment strategy of the Finnish Ministry of Education. The focus of the present study centres mainly on whether the national aims set in the framework curriculum are met equally well in all secondary schools.

The conclusions drawn from comparisons between schools have been used as indicators of equality in education [e.g. OECD: *Measuring Student Knowledge and Skills*, 2000]. A notable variation in results between different schools indicates that not all schools have been able to create equal opportunities for all children to continue their studies, or equal chances of succeeding in studies necessary for survival in a society of life-long learning.

This small study is motivated by developments made in the management of the educational system in Finland in the 1990s, according to which the creation of curricula and the allocation of educational resources has begun to be carried out locally rather than nationally. The current national framework curriculum for comprehensive schools was accepted in 1994. In terms of educational policies, the most striking characteristic of the framework curriculum is openness, as it allows schools and municipalities great freedom with regard especially to choosing content and, up to an extent, also to sharing teaching hours. Strengthening the characters of schools as well as increasing their freedom of choice were central issues discussed when the educational system was being trained to make use of the possibilities offered by the 1994 framework curriculum [Jakku-Sihvonen, 1998: 21–22]. One of the most important messages given to the local authorities was freedom and opportunity to plan educational services to fit local needs. Comparing the results achieved in each school is a way of finding out whether schools have been able to organise and implement teaching so that the aims set nationally have been met.

1.2 National strategy for assessing learning results

National assessments arranged by the National Board of Education have undergone a process of development during the last few years. The aim of these assessments is to give information on how well the aims set in the national framework curriculum have been fulfilled. In accordance with the current national strategy, a sample-based school result assessment method has been the main way of examining the effectiveness of comprehensive education.

The national assessments whose results are used in this small study were all carried out according to the same strategies, which makes comparisons possible. The series of tasks students have to complete in national assessments are created according to same kinds of principles. The tasks are created according to the current framework curriculum and levels of difficulty are set according to the criteria for final assessments given by the National Board of Education in 1999. All series of tasks have been pre-tested. Teams that are responsible for creating tasks comprise professionals in the didactics of the subject at hand, teachers who have worked in basic education and assessment professionals.

The sample includes around 5–8% of Finnish-speaking students from all secondary schools and 20% of Swedish-speaking students, from around 120 schools. The typical sample size is around 4500 students. To make sure the whole nation is represented, schools for the sample were picked by looking at the country's division into EU subsidy areas as a first criterion. Then each type of municipality (cities, suburbs and sparsely populated areas) was represented for each district in the sample. Depending on the school, either all of the school's pupils or a sample group of some students was selected.

2. EXAMINING DIFFERENCES BETWEEN SCHOOLS' RESULTS

This study is focused on finding out, in terms of educational equality, possible school-specific variations in secondary school students' learning results and attitudes, between boys and girls and between different regions in the country. This study is focused on differences between secondary schools and data is based on the national assessments. Students in the secondary schools are aged between 15 and 16, and are undergoing a transition phase leading to either vocational training or senior secondary schools.

2.1 Main question of the study

The main question is defined thus:

With regard to equality in education, are there differences between secondary schools in school-specific results and in students' attitudes towards learning?

The main question will be examined by using the following questions:

1. **Are there differences in school-specific results between schools?**
2. **Are there differences in school-specific attitudes towards learning between schools?**
3. **Are there differences between boys and girls with regard to their school-specific results?**
4. **Are there differences between boys' and girls' attitudes towards learning?**
5. **Are there differences in school-specific results between different parts of the country?**

(The concept of school-specific results is defined as the average performance of each school. School-specific attitudes are defined as the average attitude level of each school's students).

Another aim of this exercise is to find out whether it is appropriate to use a comparison between upper and lower quartile results in order to carry out meta-analytical, evaluative research.

The assumption on which the study is built is that we should expect some variation between schools' results to occur. But significant and systematic differences between schools, between boys and girls or between different operating environments are negative, as they imperil the fulfilment of students' equal opportunities for continuing studies, which is the aim of educational equality policies.

2.2 Method and building up data for meta-analysis

In order to examine differences between schools, new data were created based on school-specific results. A **new set of data was built up based on national assessments.** The following data from national assessments carried out at the end of students' secondary education were used in creating the new material:

	YEAR	Number of the schools in the original sample
Science	1998	110
Mathematics	1999	110
English	1999	125
First-language Finnish	2000	143
Mathematics	2001	112
First-language Finnish	2001	105
A-language Swedish	2001	24
B-language Swedish	2001	69

These national assessments are based on the analysis of individuals' performances in a nationally representative sample group. In all assessments a normal level of variation was observed in analysing student-specific results.

In order to carry out this empirical study, **new meta-data was created by giving each school a performance percentage (school-specific average).** This was done simply by adding up the performance percentages of each student and then dividing the sum by the number of students, to end up with an

average for the whole school. The performance percentage of a student is the score achieved by a student in the assessment out of the maximum possible score, converted into a percentage. **All schools in those eight assessments were then ranked according to their school-specific performance percentages.**

The schools in the upper and lower quartiles of the school ranking in each assessment separately were selected for this study.

Table 1. Number of schools and number of pupils in each quartile

Schools	**Sci 98**	**Maths 98**	**Maths 00**	**Eng 99**	**Fin 99**	**Fin 01**	**A-Swe 01**	**B-Swe 01**
Lower quartile	27	27	30	35	31	26	6	17
Students	803	821	1 065	1 328	1 242	853	85	661
Upper quartile	27	27	30	35	31	26	6	17
Students	840	891	1 155	1 308	1 267	969	131	787

There are 199 schools in the upper quartile and 199 in the lower quartile. Table 2 shows the number of students and the number of schools. The total number of pupils included in this study is 14,206. Only the school-specific results (averages) concerning these 399 schools were used in this exercise.

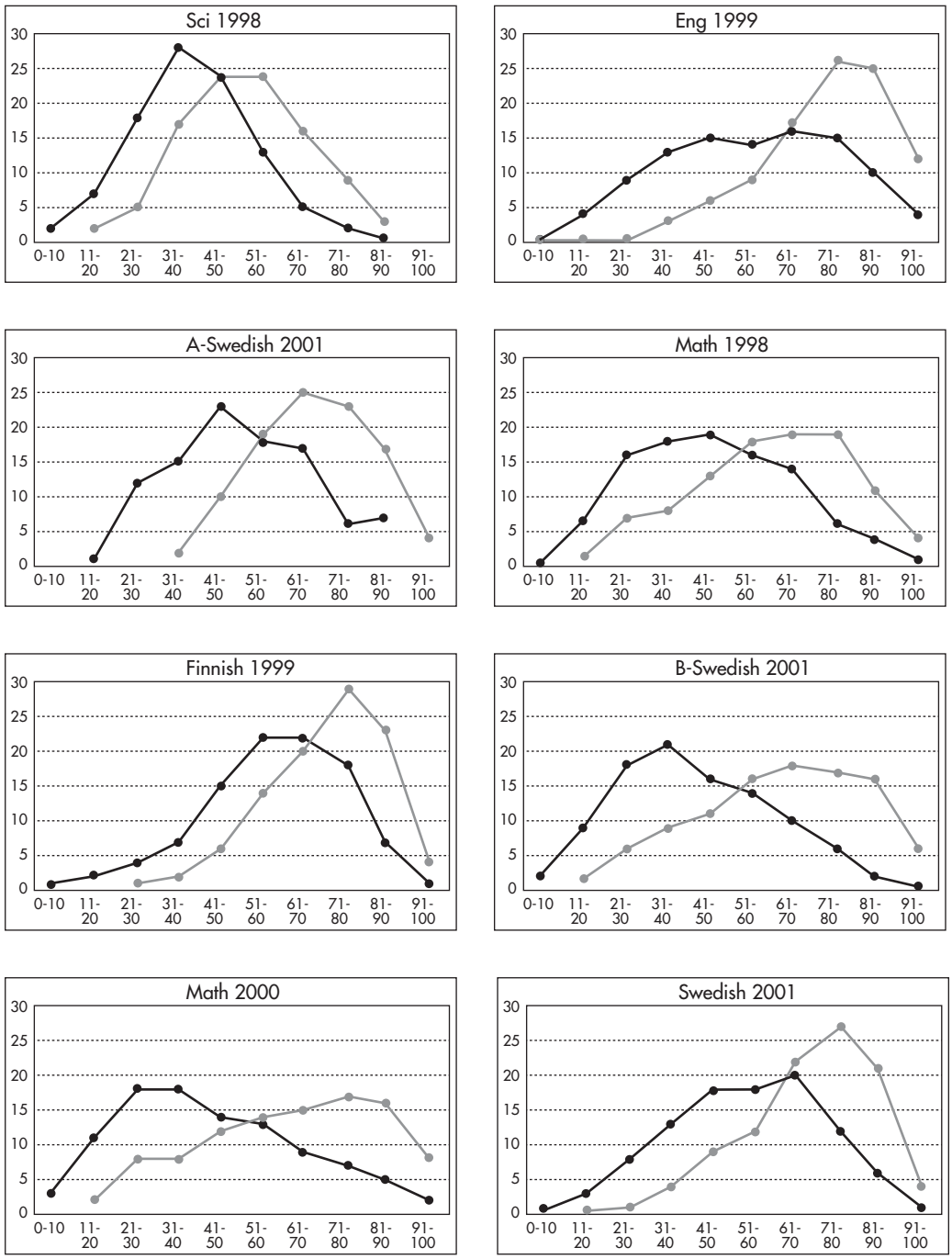
The subject-specific distributions of results, on which these data are based, are shown in Figure 1 as graphs depicting students' learning results as upper and lower quartile distributions for each subject. From the shapes of the graphs it can be deduced that in each assessment there were students in the lower quartile schools that achieved good results, in other words producing very good performances. Similarly, the data show that in the upper quartile there were schools in which a small number of students achieved a very poor level of performance. (The variation within schools is very significant, which seems to be very typical of Finnish schools. In the PISA study inside-school variation was 77% in reading literacy [Välijärvi, 2001: 27]).

Schools are figured out by different variables. Each school is defined in terms of learning by its average performance percentage. **The performance parameter is shown as a percentage of the highest possible level of performance. This variable is called the 'school-specific result.'**

The text will also used the phrases 'weaker performing schools' or 'low-performance schools' to denote the lower quartile schools, and 'better performing schools' or 'high-performance schools' to denote the upper quartile schools.

Another important variable specific to each school is the **attitude parameter**, which is a mean taken from the students' attitude scores on a scale of one to five. This is called the '**school-specific attitude.**'

The variables in this study are: school-specific learning results, school-specific attitudes towards learning, the school-specific learning results of girls, the school -specific results of boys, the school-specific attitudes of boys and the school-specific attitudes of girls, schools' quartile and schools' operating environments.



Lower quartile =black line
Upper quartile = gray line

Figure 1. The individual performance percentages of the upper and lower quartile schools' students in different assessments.

Schools' operating environments refer the part of the country where the schools are located: capital city area, northern Finland (regions of Lapland and Oulu) or the rest of the country.

The data was analysed mainly by comparisons (t-test, variance analysis and Chi2-test). In this report the way of comparing schools is very rough. We simply compare the schools-specific averages (whole school, boys' average, girls' average) and we compare the number of the lower and upper quartiles schools in different parts of the country. The former comparison is for finding out if there are differences between schools or between boys and girls, and the latter is for examining if there are differences between regions in terms of the number of the lower-performance schools and the high-performance schools.

Results are reported in detail, by question, in order to find evidence for evaluating equal opportunities in education in terms of the learning results, attitudes and operating environments of the schools.

This study is very much an exercise for finding out a very quick and rough method for evaluating politically interesting trends.

3. VARIATIONS BETWEEN SCHOOLS' RESULTS

At the basis of the research is the assumption that in basic education, variation exists *within* the results of individual schools, but not significantly *between* schools. Assessments carried out up to now have found that school-specific variation does, however, exist at least in mathematics and first-language Finnish results [Korhonen, 2001; Lappalainen, 1999 and 2001].

The PISA study (Programme for International Student Assessment) carried out by the OECD produced information concerning differences between schools, based on what percentage of the total variation in a student's performance is attributable to variation between schools. In the case of Finland the figure is 10.7%; on average, across the OECD countries, it is 36 per cent of the total variation [PISA report, 2001: 197, 257].

According to Välijärvi, in reading, only five percentage points of variation in results can be attributed to the school [Välijärvi, 2001: 27]. The PISA study, on the other hand, found that a school's effect on a child's performance can account for up to 10 percentage points for boys and four percentage points for girls [Välijärvi *et al.*, 2002: 29]. Hautamäki argues that a school's effect on results could even be as significant as 20 percentage points [Hautamäki, 2000: 251].

The method used in this study is very rough and totally different from those presented above. Only simple comparisons between schools are used.

This exercise parts from the premise that if there is a difference of over 10 percentage points in the average performance levels of schools in the upper quartile and schools in the lower quartile, we should consider this to be evidence of the existence of differences between schools, significant in terms of the levels of equality they can offer students. The choice of the 10 percentage point figure is based mainly on a personal conclusion, which stems from the fact that

if the difference in schools' averages is of over 10 percentage points, this could signify a difference of over one grade (on a scale of four to ten) in the average grade of the students' school-leaving reports. That big a difference in average grades may be significant in student's opportunities for continuing their studies in very popular trades or in senior secondary schools, as the school-leaving report is the main criterion used in selection.

The significance of the difference is marked in its original form with asterisks, such that

p = 0.05 = *	df = degree of freedom
p = 0.01 = **	N = Number of schools
p = 0.001 = ***	p.p.= percentage point (unit)

3.1 School-specific differences in learning results

Within the lower quartile school-specific averages vary between 31% and 61% depending on the subject of the assessment and in contrast, within the high-performance schools between 48% and 87%. (Because the data have been selected by taking the highest and lowest quartiles out of the results of each assessment, the percentages overlap to a certain extent. This is due to the fact that in some assessments the average performance level have been lower than in others.)

In the school-specific meta-data, depending on the subject, the average of the schools-specific performance percentages in the upper quartile schools ranges between 52% and 74%, whereas in the lower quartile schools it ranges between 39% and 58%. In all the assessments the smallest difference between the averages of the upper and lower quartile schools' performance percentages is of 13 p.p., and the largest difference is of 21p.p. (Figure 2).

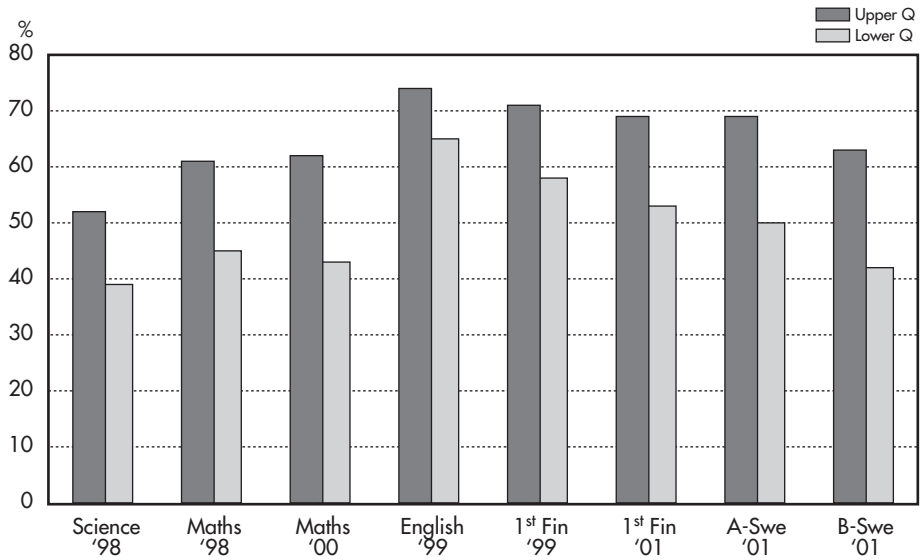


Figure 2. Average school-specific results in the upper quartile and lower quartile.

It is interesting that the smallest difference between quartiles is in Finnish language. Also in international comparisons the level of skills in reading has been found to be very high inFinland [cf. the PISA study].

Looking at all the subjects together as a whole, within the lower quartile schools the average performance percentage is 49 (with a deviation of eight p.p.) and within the upper quartile schools it is 65 (deviation of eight p.p.). The difference is 16%.

With regard to question 1, we can conclude that the difference between the averages of upper and lower quartile schools in the meta-data, when compared on a subject-specific basis, ranges between 13 percentage points and 21 percentage points, and the average difference between their average performance percentages is of 16 percentage points. This difference converted into actual numerical grades could mean a deviation of one to two grades out of ten in each school's average. A difference as significant as this, if it indeed does become visible in students' school-leaving reports, can mean an important imbalance in students' opportunities for continuing studies after finishing their basic education. The difference presumably also shows an imbalance in students' actual skills when finishing their basic education, and their actual capabilities for continuing studies.

3.2 School-specific attitudes towards learning

The examination of students' attitudes towards learning in national assessments takes into account amongst other things whether students see a subject as significant, whether they enjoy the subject, whether they have experience in

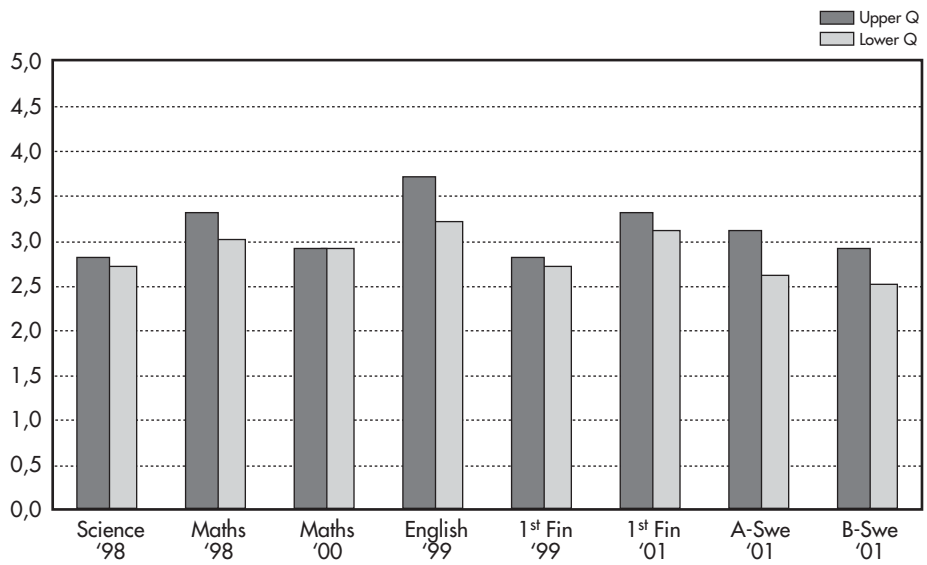


Figure 3. Average attitude levels within the different subjects in upper and lower quartile schools.

succeeding in the subject and how they relate to studying the subject. In assessing attitudes on an individual basis, it has been noted that on the whole, girls' attitudes are more positive than boys', and that successful students' attitudes are more positive than less successful students'. On the basis of this, we can expect attitudes between the high-performance schools and the low-performance schools also to vary.

The upper quartile schools' average attitude level scores ranged between 2.8 and 3.7, and the lower quartile schools' average scores ranged between 2.5 and 3.2 (Figure 3).

The average attitude level in the lower quartile schools was 2.9 (deviation of 0.29) and in the upper quartile schools it was 3.1 (deviation of 0.36). A variance analysis shows that there is a statistically significant difference between the attitudes of students in upper quartile and lower quartile schools (sign. ***).

Concerning question 2, these results mean in practice that psychological learning atmospheres may vary between high- and low-performance schools.

3.3 Differences between girls and boys with regard to their school-specific results

In all assessments except the mathematics and science results showed girls performing better than boys. Even in biology and geography girls came out better than boys. The international PISA study also found gender-based differences in performance. For instance in reading literacy, girls were clearly above boys. According to the PISA study the magnitude of the difference between girls' and boys' performances was largest in the lower quartile schools and smallest in the upper quartile schools [Välijärvi, 2001: 29].

3.3.1 School-specific differences between girls' and boys' learning results

Based on the meta-data, the difference between boys' and girls' school-specific results is statistically significant when all the subjects are taken into account (sign***).

The girls' average is 60%, with a deviation of 12, and the boys' average is 54%, with a deviation of 13. The boys' average in the upper quartile schools was 63%, with a deviation of eight, and in the lower quartiles 45%, with a deviation of eight. The average performance of girls is in the upper quartile was 67%, with deviation of 11, and in the lower quartile 53%, with deviation of 11.

Within each quartile, the school-specific averages of boys and girls also differ from each other in a statistically significant way (sign.***).

The conclusion for question 3 is that if all subjects are taken into account, girls perform better than boys in both the upper and the lower quartile schools. In the upper quartile schools the difference between girls and boys is not

nearly as large as in the lower quartile schools. (When examining the matter by subject we find that upper quartile girls are notably better than boys in the language subjects, but upper quartile boys are as good or better than girls in mathematics and the sciences).

For the sake of comparison we may mention here relevant results from the PISA project. When looking at differences between girls and boys from the results of that study, it was found that within one school the numerical difference between girls' and boys' averages was 126 points (the average grade for the test internationally was 500 points). When in the same data schools were separated into quarters or quartiles, girls' performances in the lower quartile schools were on average 66 points higher than boys'. Also in the upper quartile schools girls came out clearly better than boys, with 38 points more on average [Välijärvi *et al.*, 2001].

3.3.2 Differences between girls' and boys' schools-specific attitudes

On the basis of school-specific averages there is a statistically significant **difference in attitude between girls and boys; girls relate in a more positive way** to the study of the assessed subjects and to the usefulness of the subjects than boys. The girls' average attitude score was 3.04, with a deviation of 0.57, and the boys' was 2.92, with a deviation of 0.44 (sign.**).

It is, however, remarkable thatdifferences exist between quartiles. Only in the lower quartile schools was there a statistically significant difference between girls and boys such that girls felt more positively towards studying than boys (sign.***).

In the upper quartile there was no difference between girls' and boys' attitudes towards studying.

The average of girls' attitudes in the lower quartile schools was 2.94 (deviation of 0.51) and the boys' was 2.79 (deviation of 0.39). In the upper quartile schools, the girls' average was 3.14 (0.60) and the boys' was 3.04 (0.47).

3. 4 The upper and lower quartiles schools in different parts of the country

The operating environment was selected as one of the main focal points of the study because other analyses of national assessment data have found there to be connections between the type of municipality, the EU subsidy area and the students' performances [cf. e.g. Tuokko, 2001; Rajakorpi, 1999; Lappalainen, 2000 and 2001; Toropainen, 2002]. Also the results of the PISA project indicated that students' results appear to be slightly weaker in rural environments than in urban areas. Other geographical differences apparent from the PISA study were very small and thus not statistically significant. When results were examined from the point of view of geographical areas and socio-economic

backgrounds, differences were, however, found. This indicates a correlation of some sort between school results, population groups and living environments [cf. Välijärvi *et al.*, 2001: 30 and 2002: 110–112].

In order to analyse schools' operating environments from a regional point of view, we created an area classification system. In this classification the areas' levels of education and employment were very important identifying factors. (In the last few years unemployment levels have reached 7% in the capital city areas, 11% in the whole country and around 16% in northern Finland). The data were classified so that the capital city area (Helsinki, Espoo and Vantaa) makes up one category, northern Finland (the districts of Oulu and Lapland) makes up the second category and the rest of Finland forms the third category.

3.4.1 School-specific' learning results in different parts of the country

In order to analyse the differences between school-specific results in different parts of the country, the data were cleaned so that no school was included more than once. Table 2 brings together the performance averages and numbers of schools sorted by geographical location in the country. It can be observed that performance levels between quartiles vary a great deal in the capital city area. The average performance level of the upper quartile schools improves as we move south from northern Finland, and the average performance level of the lower quartile schools falls.

The proportion of high-performance schools in the capital city area is larger than that of the low-performance schools. In the northern part of Finland the opposite is true, and the proportion of low- performance schools is higher than that of high-performance schools. In the rest of Finland the situation is even, with almost the same number of low- and high-performance schools.

Figure 4. shows the distribution of schools.

When examining the distribution of upper and lower quartile schools in various parts of the country, we can see that the division of schools into quartiles differs significantly from expected values (CFA by Alexander von Eye), (sign.***).In this analysis the number of the upper level schools in the northern part of the country is significantly lower than expected.

Table 2. Distribution of schools in the capital city area, in northern Finland and in the rest of Finland, and their average performance level percentages

Area	Lower quartile schools		Upper quartile schools	
	Performance level %	N (%)	Performance level %	N (%)
Capital city area	46	18 (31)	69	41 (69)
Other districts	49	127 (50)	64	126 (50)
Northern Finland	49	39 (75)	63	13 (25)

N = number of schools

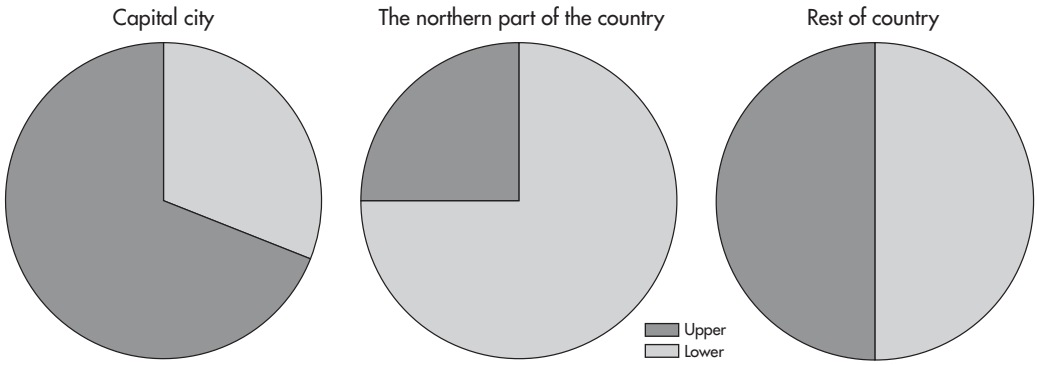


Figure 4. The distribution of schools.

Table 2 shows an important factor in terms of regional equality: the capital city area is exceptional in terms of the polarisation. In practice this means that the area's best schools are really good and its worst schools include some really poor ones. The performance level average in the capital city area's high-performance schools (69%) is the highest in the country, and the average in its low-performance schools (46%) is the lowest.

The performance level in the capital city area is higher than in the northern part of the country (sign ***) and in the rest of the country (sign ***). If all subjects are taken into account, the difference between the Northern part of the country and the capital city is 9 percentage points.

On the basis of these analyses we can reasonably observe that there are some differences between the average learning results of schools located in different parts of the country. In the northern regions there are less high-performance schools as in the capital city area. It is worth of noticing the risk of polarisation that exists in the capital city. The schools in the upper quartile have achieved very good performances, whereas the lower quartile includes schools that achieve very low results.

When looking at the differences in girls' and boys' performances in different operating environments (considering all subjects at the same time) we can observe that there are fairly clear differences between performances in schools within the capital city area. Both the worst and best averages of both boys' and

Table 3. School-specific performance percentages and deviation in different parts of the country

Area	N	Average performance %	Deviation
Capital city area	68	62	13
Other districts	275	57	11
Districts of Oulu and Lapland	55	53	9

N = Number of schools

Table 4. Performance and deviation by quartile, gender and the part of the country

Area	Lower quartile				Upper quartile			
	Boys		Girls		Boys		Girls	
	Average	Dev.	Average	Dev.	Average	Dev.	Average	Dev.
Capital city area	44	10	49	12	68	9	68	10
Rest of Finland	46	8	53	10	62	8	66	11
Districts of Oulu and Lapland	45	7	54	11	58	7	64	12

girls' school-specific averages can be found in this area (cf. Table 4). **In terms of equality between genders it is noteworthy that in the capital city area there are no gender-specific differences apparent in upper quartile schools (Figure 5).**

In northern Finland and the rest of Finland apart from the capital city area, the difference between boys' and girls' performances is clear: girls perform better than boys. In the northern part of the country the difference between girls' and boys' performances can be of up to nine percentage points, whereas in the capital city area it only reaches five percentage points. It is interesting to note that in northern Finland the difference between the girls in the worst-performing schools and the boys in the better-performing schools is just four percentage points, whereas in the capital city it is 19 percentage points, and in the rest of Finland nine percentage points. **The difference in the performances of boys in the capital city's lower and upper quartile schools is 24 percentage points**, which could mean a difference of over two numerical grades when converted into a school's average performance grade. This figure in northern Finland is only 13 percentage points or about one grade.

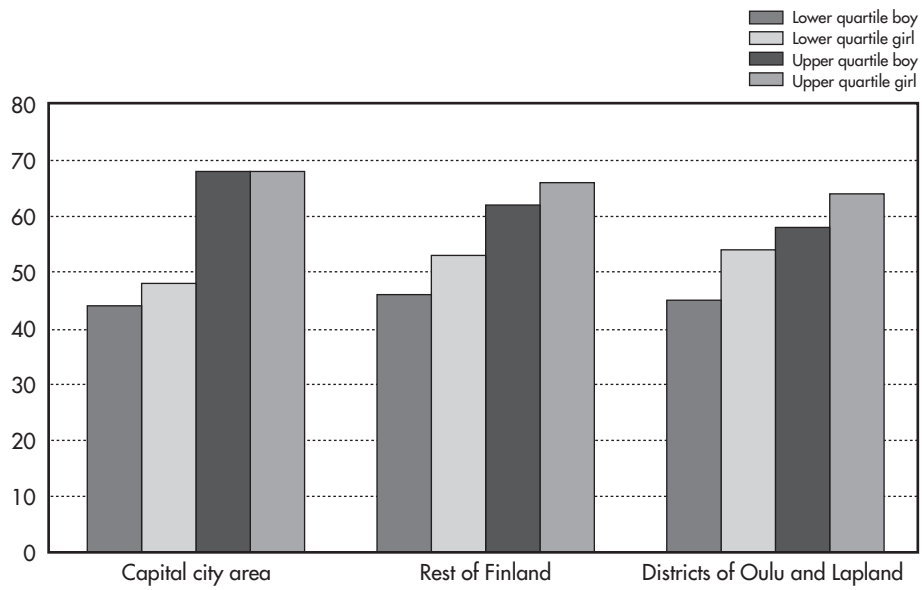


Figure 5. School-specific performance levels for boys and girls by quartile and part of the country (including all subjects).

The conclusion for question 5 is that a school's operating environment is linked to the school's results. The observations regarding the capital city area, according to which girls and boys achieve equally good averages are very important and encouraging. Based on this study it appears that the variations in the school-specific performances of girls and boys in different operating environments is telling us something important about the contrasting expectations that we hold of boys and girls in terms of their socio-cultural environments and studies.

3.4.2 Do differences between schools exist in terms of equality in education?

The main question of the study was formulated as follows:

With regard to equality in education, are there differences between secondary schools in school-specific results and in students' attitudes towards learning?

The motivation for this study arises from the definition that in terms of equality in education it is important for a school's operating environment not to have any effect on the school's learning results or attitude . For the sake of achieving equality in education, schools should offer each student an equal opportunity of achieving good results at national assessments, regardless of their operating environments. Equality between boys and girls means that no systematic differences exist.

On the basis of the results of the various questions the main problem was broken down as, we can reasonably conclude that some differences do exist between schools in terms of educational equality, when school-specific learning results are used as a criteri on. School-specific learning results tend to vary, to some extent systematically between regions if the number of the upper and the lower quartile schools in different regions is used as a criterion for this matter. Equality between boys and girls seems to be more real in the better performing schools of the capital city area than elsewhere. Attitudes tend to vary between quartiles.

The main findings are as follows:

1. **The differences in school-specific averages between quartiles ranged between 13 and 21 percentage points.** These differences in school-specific learning averages signify a difference of one to two numerical grades.

2. There are also differences in students' attitudes between schools, such that **attitudes towards studying tend to be more positive in the schools that perform better.**

3. There are differences in the learning results achieved by boys and girls in weaker-performing schools : girls achieve better results than boys. **In better-performing schools differences between boys' and girls' performances are smaller or non-existent.**

4. Differences between the lower and upper quartile are marked in all parts of the country, but they as very remarkable in the capital city area.

5. **Northern Finland has proportionally the lowest number of high-performance schools.**

In a report included in this publication, Dr. Jorma Kuusela examines the connections between learning results and demographic factors.

4 VALIDITY OF THE RESULTS

This study was a new kind of attempt at carrying out a meta-analysis of assessments that measure national schools achievements. Main motive was to find out trends in educational equality.

The aim of the methodology used was to compress information received from the national assessments. The validity of the data available is enhanced by the fact that the assessments were carried out on several different subjects and that they dated from different years. The assessment system has purposely been structured so that data is gathered each year using the same systematic methods. This is why it has been possible to create new meta-data to suit our purposes.

Especially the results concerning attitudes are very valid based on the fact that all of them have been assessed by using same scale. Concerning the school-specific learning results there are some risks that has to be taken into account. The most problematic risk is that in different assessments the national level has varied. That is why we made this small exercise by using only very rough comparisons for finding out differences. We only compared the average percentages between quartiles. Variance analysis method was used because of the data very carefully and only when the quality of the data made it possible without any big risks. In case of the regional analysis only the number of schools in different regions were compared.

The study was designed to focus on the best and worst performing schools for finding out politically important trends. The basis for this choice was that in terms of achieving equality in education, it is the size of the differences and the systematic differences between schools that are important for policy-makers. This makes the comparison of the lower and the upper quartiles suitable. By using data gathered in several assessments it is possible to work out whether the differences are systematic and meaningful.

The most problematic issue is that due to the way in which information was gathered, some schools have taken part in more than one assessment, which could have had a slight effect on the results. The number of these schools in

the data is so limited, though, that it should not have such an effect. Still, this was taken into account when comparing the school-specific results of different regions. The data were cleaned so that no school was included twice in the analysis.

The numbers of girls and boys in the various sample groups vary, but that should not have a significant impact, either, on conclusions drawn from such a large base of material.

In evaluating the validity of this meta-analysis it is important to note that when examining an assessment study, its relevance to society and the usefulness of its conclusions are very significant factors in its worth [House, 1980: 249–250; Patton, 1997: 122; Jakku-Sihvonen, 2001: 132; Huhta & Takala, 1999: 213].

5 DISCUSSION

In order to promote equality in education, it is crucial that during basic education we strive to prevent the appearance of differences in learning results and attitudes. This is because a student's basic education opens up the student's path into higher education and forms a basis for his or her lifelong learning.

Factors related to schools' operating environments appear to be closely related to the environment's education level and socio-demographic characteristics [this question is analysed further on the basis of the same data by Jorma Kuusela, in this publication.]

It would be especially important to work out the psychological and social characteristics of each operating environment. The polarisation apparent between schools in the capital city area gives cause to further research.

The systematic nature of differences between girls and boys in the lower quartile schools is worrying. Does this point towards the fact that there are differences in the environment's expectations of girls' and boys' efforts and ways of working with regard to schoolwork? Are boys allowed more freedom to choose which rules and responsibilities are worth following? It is important to work out how the best-performing schools have managed to create a learning environment that is supportive and induces good achievements in both girls and boys.

Policymakers need evaluation for finding out warning signs as well as information concerning positive development. Evaluators have to find out new ways of finding out information worth of intensive studies. The comparison of upper and lower quartiles appears to give useful information for an evaluator concerning positive and negative signs. Based on this study, we would be justified in believing that methods based on comparing upper and lower quartiles can be developed as a method of evaluation, in order to find warning signals or positive signs worth of more careful research. Using material drawn from the upper and lower quartiles it is possible to compress a large amount of data so that we are able to tackle the question with the help of simple indicators.

BIBLIOGRAPHY

Hannén, K., 2001: *Utvärdering av inlärningsresultat i modersmål i åk 6 i den grundläggande utbildningen våren 2000*. Finnish National Board of Education. Utvärdering av inlärningsresultat 4/2001. Helsinki.

Hautamäki, J., Arinen, P., Bergholm, B., Hautamäki, A., Kupiainen, S., Kuusela, J., Lehto, J., Niemivirta, M. & Scheinin, P: Oppimaan oppiminen ala-asteilla. *Finnish National Board of Education. Oppimistulosten arviointi 3/1999. Helsinki.*

Hautamäki, J., Arinen, P., Hautamäki, A., Ikonen-Varila, M., Kupiainen, S., Lindblom, B., Niemivirta, M., Rantanen, P., Ruuth, M. & Scheinin, P: Oppimaan oppiminen yläasteella. *Finnish National Board of Education . Oppimistulosten arviointi 7/2000 . Helsinki.*

House, E., 1980: Evaluating with Validity. *Bevery Hills, CA: Sage.*

Huhta, A. & Takala, S., 1999: "Kielitaidon arviointi", within the work by Sajavaara, K. & Piirainen-Marsh, A. (eds.): *Kielenoppimisen kysymyksiä. Soveltavan kielentutkimuksen teoriaa ja käytäntöä.* University of Jyväskylä.

Jakku-Sihvonen, R., 1996: "Opetuksen saavutettavuus ja koulutuksellinen tasa-arvo arvioinnin kohteena", within the work *Toteuttaako peruskoulu tasa-arvoa?* Finnish National Board of Education. Arviointi 1/1996. Helsinki.

Jakku-Sihvonen, R., 1998: "Koulu uudistusten paineissa". In the report: Niemi. H. (ed.): Opettaja modernin murroksessa. *Opetus 2000.Jyväskyä: PS-viestintä Oy.*

Jakku-Sihvonen, R., 2001: "Arvioinnin luotettavuudesta", in the work by Jakku-Sihvonen, R. & Heinonen, S.: Johdatus koulutuksen uudistuvaan arviointi-kulttuuriin. *Finnish National Board of Education. Arviointi 2/ 2001. Helsinki.*

Korhonen, H., 1999: *Peruskoulun matematiikan oppimistulosten kansallinen arviointi 1998*. Finnish National Board of Education. Oppimistulosten arviointi 1/1999. Helsinki.

Korhonen, H., 2001: *Perusopetuksen päättövaiheen matematiikan oppimistulosten kansallinen arviointi 2000*. Finnish National Board of Education. Oppimis-tulosten arviointi 3/2001. Helsinki.

Finnish National Board of Education, 1998: "Koulutuksen tuloksellisuuden arviointimalli." Arviointi 7/1998. Helsinki.

Kupari, P., Reinikainen P., Nevanpää, T. & Törnroos, J., 2000: *Matematiikan ja luonnontieteiden osaaminen peruskouluissa* (summary). Department of Pedagogical Research, University of Jyväskylä. Jyväskylä.

Lappalainen, H-P., 2000: *Peruskoulun äidinkielen oppimistulosten kansallinen arviointi 9. vuosiluokalla 1999*. Finnish National Board of Education. Oppimistulosten arviointi 1/2000. Helsinki.

Lappalainen, H-P., 2001: Perusopetuksen äidinkielen ja kirjallisuuden oppimis-tulosten kansallinen arviointi 9. vuosiluokalla 2001. *Finnish National Board of Education. Oppimistulosten arviointi 6/2001. Helsinki.*

Metsämuuronen, J., 2001: "Monimuuttujamenetelmien perusteen SPSS-ympä-ristössä." *Metodologia* series 7. International Methelp Ky. Estonia. Pp. 104–114.

Numminen, J., 1994: *Koulutuspolitiikan vaihtoehdot*. Tampere :Tammerpaino.

OECD, 2000: ”Measuring Student Knowledge and Skills”. *The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*. Education and Skills.Paris: OECD

Patton, M., 1997: Utilization Focused Evaluation. *The New Century Text. Thousand Oaks, CA: Sage.*

Rajakorpi, A. (ed.), 1999: *Peruskoulun 9.-luokkalaisten luonnontieteiden oppimistulosten arviointi. Keväällä 1998 pidetyn kokeen tulokset.* Finnish National Board of Education. Oppimistulosten arviointi 2/1999. Helsinki.

Rinne., R., 2001: ”Koulutuspolitiikan käänne ja nuorten syrjäytyminen”, in the report: Jauhiainen, A., Rinne, R. & Tähtinen, J.: *Koulutuspolitiikka Suomessa ja ylikansalliset mallit*. Suomen kasvatustieteellinen seura 2001. Turku.

Toropainen, O., 2002: *Nationell utvärdering av inlärningsresultat i finska i åk 9 våren 2001.* Utbildningsstyrelsen. Utvärdering av inlärningsresultat 1/2002. Helsinki.

Tuokko, E., 2000: *Peruskoulun 9. vuosiluokan englannin (A1-kieli) oppimistulosten kansallinen arviointi 1999.* Finnish National Board of Education. Oppimistulosten arviointi 3/2000. Helsinki.

Tuokko, E., 2002: *Perusopetuksen päättövaiheen ruotsin kielen oppimistulosten kansallinen arviointi 2001.* Finnish National Board of Education. Oppimistulosten arviointi 3/2002. Helsinki.

Vasama, P. M. & Varia, Y., 1972: *Johdatus tilastotieteeseen*. Helsinki: Yliopiston tuki ry.

Välijärvi, J., Linnankylä, P., Kupari, P., Reinikainen, P., Malin, A. & Puhakka, E., 2001: *Suomen tulevaisuuden osaajat*. Department of Pedagogical Research, University of Jyväskylä, Finnish National Board of Education & OECD PISA.Jyväskylä.

Välijärvi, J. & Linnankylä, P. (eds.), 2002: *Tulevaisuuden osaajat*. PISA 2000 in Finland. Department of Pedagogical Research, University of Jyväskylä, Finnish National Board of Education & OECD PISA. Jyväskylä.

Wuolijoki, H., 1999: *Matematiikan oppimistulokset ammatillisessa perustutkinnossa*. Finnish National Board of Education. Oppimistulosten arviointi 5/1999. Helsinki.

Jorma Kuusela

LINKS BETWEEN SCHOOL RESULTS AND DEMOGRAPHIC FACTORS

1 AIM OF THE STUDY

The aim of this article is to describe the connections that exist between students' school results and factors related to demographic structures.

In order really to understand students' learning-based results, we need to obtain a more complete vision of students' performances than we would receive by simply looking at individual subjects. Therefore, to build a more complete picture, in this study we have gathered together data from various different sources. Even though different things tend to happen with different subjects, it is difficult to justify the use of single subjects in isolation. The issue may be to do with real subject-specific differences or simply with differences in the ways assessments have been focused; even with unusual things happening in the selected samples. All the subjects selected for use in this study are ones that are central to basic education (in Finland the combination of primary and secondary schools, teaching children between the ages of 7 and 16) and ones which have recently attracted special attention in public education-related discussions. It is also obvious that the assessment programme created by the National Board of Education stems just from these central subjects.

As the aim of the study is set to be the definition of points of connection between learning-based results and demographic factors, we do not necessarily have to formulate a starting hypothesis. A natural presumption is that these points of connection exist. For instance Professor Kuusinen indicates in his studies that from the point of view of the student, the links between a student's learning-based results and educational choices and the parents educational background are clear [c.f. for instance the work of Nummenmaa *et al.*, 1997]. It has also been proved that parents' educational backgrounds are linked to students' skills in learning to learn [Hautamäki *et al.*, 1999, 2000, 2002]. On the other hand, the latest report on the PISA project emphasises the regional uniformity of the learning-based results obtained [Välijärvi *et al.*, 2002]. It is also possible that the statistical data currently available will not demonstrate the kinds of connections we are looking for; assessments carried out so far for instance by the National Board of Education have not found particularly large differences in students' results beween different areas (districts, municipalities classified statistically, EU objective programme areas), and if differences have been apparent they have varied from one assessment to the next. The meta-analysis carried out by Ritva Jakku-Sihvonen, whose results are described in this very publication, indicates, however, that regional differences exist which do not become apparent simply by comparing districts.

If there are links between school results and demographic factors, it is important to find them for at least three good reasons. Firstly, it should be integral to our interest in the matter that we cannot simply assume that school results are completely independent of students' living environments. Secondly, if results are indeed related to demographic factors, measures may have to be taken with regard to this, especially because demographic differences appear to be increasing rather than decreasing in Finland. It is difficult to prevent or slow down the appearance of differences between regions, but it would be easier to try to reduce the impact of these differences on students' results, at least if the trends are noticed early enough. Currently we can still assume that these are early stages, because if the regional differences between students' results were obvious and alarming, they would have been noticed and reported on long ago.

The third reason is related to assessments. According to new education laws in Finland, assessment is an important development tool, which is at least partly seen as replacing norm and resource-based management. Assessments must be used by the national educational system and regional educational governments as well as individual schools. Data simply about whether single schools' results are good, average or poor cannot be used as the best kind of development tool. More analytical information is needed. Assessment must also be fair in the sense that it should include and show the links between any differences in teaching principles or schools' operating principles and students' final results. The effectiveness of schools can then be examined from the point of view of those kinds of background elements. When we talk about the added value of schools we mean the ways in which schools differ from what is predicted on the basis of various background variables. Increasing that added value is probably one of the greatest challenges in education.

Our final definition with relation to this study is to do with the term 'connection' or 'link.' Our aim at this point is not to find causes and consequences; this kind of material does not allow us to do that. In speaking of connections or links, we are interpreting students' learning-based results as one of many regionally descriptive indicators. Let us emphasise as well that these indicators describe exactly that: the area, not individual students. Perhaps the most common false deduction that these kinds of studies tempt people to make is known as an 'ecological error,' 'wrong level error' or 'Durkheimian error.' These terms signify the mistake of making micro-level inferences from macro-level data: in other words, regional trends are generalised to relate to individual people or families. The most famous example of this is probably a study made in the US, which found that crime levels were highest in the areas inhabited by blacks. The prejudice-based conclusion was drawn that black people commit more crimes, until more detailed research showed that the crimes were actually being committed by the white people who for one reason or another had ended up living in the predominantly black-inhabited low status areas. The fact that the area was inhabited mainly by black people was a good regional indicator, but the generalising conclusion drawn from the results was false. Demographic links are often complex, and already because of that it is good to avoid speculation on causes and consequences.

2 MATERIALS

By 'students' results' or 'learning-based results' we signify in this study the national result assessments carried out on the final stages of basic education between 1998 and 2001 by the National Board of Education. These assessments were carried out on samples, with 5-7% of the students in the age group taking part in each assessment. In all cases the samples have been based on random distributed cluster sampling. The points of view from which we have looked at the equality of the distribution are districts, the statistics-based division into municipalities (rural, suburban and urban municipalities) and EU subsidy areas (nowadays called EU development areas). Monitoring the distribution of the set of samples has ensured that each assessment gives a reliable overview of the situation in the whole of Finland, and the divisions into areas mentioned above are considered essential in terms of regional equality. By cluster sampling we mean that the sample units are whole schools and not individual students.In big schools students have been picked through an even spacing sample technique. In using these techniques we lose the opportunity to make generalisations that is a part of pure random sampling, but what we chose makes more sense in practice. By randomness in our method we mean that within each area, each school has an equal chance of being picked.

All result assessments used here as material were made on the final stage of basic education, in other words in the ninth grade (students aged 15–16). The assessed subjects are:

- Mathematics [Korhonen 1999]
- Science [Rajakorpi 1998]
- Finnish as a first language [Lappalainen 2000]
- English (as an A1 language) [Tuokko 2000]
- Mathematics [Korhonen 2001]
- Finnish as a first language [Lappalainen 2001]
- Swedish as a second language [Tuokko 2002]

For Swedish we assessed in this study students who had followed the course Swedish as a B language, the course that teaches the minimum level that is compulsory for all students. Students may choose to take the optional, more advanced course of Swedish as an A language instead. We did not choose to use results from assessments of A language students, because the teaching and selection of that course is related to kinds of background variables that are not suitable for this analysis. Most often A-language Swedish is only offered in larger towns and bigger schools, and generally it is the more motivated students, or at least the children of more motivated families, who choose to take it. Around two thirds of the students in A-language Swedish are girls. If we chose to use A-language Swedish, which occurs only in a few areas, and treated it as equivalent to other, more widespread subjects, it could distort our results.

In the analyses we have used the actual original data rather than reports on it.

Table 1. Assessments used in this study

Assessment	Year of assessment	Municipalities	Schools	Students
Mathematics	1998	92	110	3575
Science	1998	85	109	3395
Finnish as a first language	1999	89	125	5137
English (as an A1 language)	1999	103	143	5641
Mathematics	2000	87	112	4129
Finnish as a first language	2001	81	105	3891
Swedish (as a B language)	2001	60	69	2869
Total number of students assessed				28637

In total 291 municipalities, in other words nearly two thirds of the municipalities in Finland, took part in the assessments dealt with here, and the total number of students assessed is equal nearly to one half of the students in that age group in Finland. There were schools from all the biggest cities taking part in each assessment. 12 municipalities took part in at least six assessments, 19 in at least five, 36 in at least four, 69 in at least three and 162 in at least two. How many assessments each municipality takes part in depends mainly on the size of the municipality, or, to be precise, on the number of schools in it. Another factor is the combination of variables that applies to each municipality. A municipality that belongs to an unusual combination of variables has a better chance than others of ending up in the sample group more than once.

Population data have been acquired from two sources: the Kuntafakta (Municipality Fact) database created by Statistics Finland, and the study by Ville Viljanen ”Huono-osaisuuden alueellinen kehitys 1990-luvulla” (”The regional development of deprivation in the 1990s,” Association of Finnish Local and Regional Governments, 2001. Acta Series, publication 139).

3 APPROACH AND RESEARCH STRATEGY

The basic unit on which analysis was based in this study is the municipality, as the demographic data available is municipality-based. Average municipality-based data, however, do not, especially in the larger areas, describe without fault the demographic circumstances of the areas which students are actually from (rather than the areas where they attend school). Until now most analyses have been carried out with the variance analysis method, comparing the variables mentioned above, and the basic unit for the analysis has been the student. The regions into which the country is usually divided are so large that differences are very small; when all regions are treated as homogeneous for the sake of analysis, internal variation is not taken into account, and therefore results in larger regions will be more inaccurate. When this is combined with large sample numbers, even small variations come out as statistically significant, and at worst significance-based studies end up being simply reports on trivialities. Another way of looking at differences in the past has been the study of ‘explanatory proportions.’ As the measurement of an explanatory proportion the studies

have used the square of the correlation or regression coefficient, or the square of the Eta, which measure up to what extent differences between students can be explained according to the differences between the background factors. If the primary unit for the analysis is the student, explanatory proportions remain very small. In these cases researchers have ended up in paradoxical situations in which the links found are 'very significant,' but don't actually 'explain' anything.

If the links between students' results and their regional background variables were looked at as simple explanatory proportions it would be easy to forget how large the differences between individual students are in reality. Around ten per cent of of the differences between students' results can, however, be interpreted as variations between schools, which is not at all a small proportion.

Table 2. The portions attributable to schools out of the total variation in students' learning-based results

Assessment	Year of assessment	Proportion attributable to schools out of total variation
Mathematics	1998	8 %
Science	1998	11 %
Finnish as a first language	1999	9 %
English	1999	11 %
Mathematics	2000	11 %
Finnish as a first language	2001	11 %
Swedish	2001	15 %

If we look at the regional division-based variables the explanatory proportions naturally decrease significantly.

Table 3. Portions attributable to the variables out of the total variation

		Portion out of total variation		
Assessment	Year of assessment	District	Municipality	EU subsidy area
Mathematics	1998	0.2%	0.0%	0.5%
Science	1998	0.9%	0.3%	0.7%
Finnish as a first language	1999	1.1%	0.2%	0.4%
English	1999	2.2%	1.1%	3.6%
Mathematics	2000	0.3%	0.3%	0.3%
Finnish as a first language	2001	1.2%	0.2%	0.4%
Swedish	2001	1.4%	2.2%	1.9%

The effect of regional factors should really be visible in average results rather than individual students' performances. This is why we should choose as the primary unit for the analysis either the school or the municipality. In doing this we would be assuming that differences between larger units such as these cannot be considered to be insignificant. In this study we will examine differences between municipalities. Each municipality in Finland is represented in the study by one or more schools, which have each taken part in one or more assessment. The selection of the municipality as the unit to be analysed is justified because it is possible to obtain demographic data regarding municipalities from public databases. The acquisition of school-specific data would require creating a whole separate database. At least at this point, economy takes precedence over accuracy.

The organisation of this study cannot be considered entirely flawless, because differences internal to municipalities may still exist, which will not become apparent in this study (due to the types of data available). This can happen especially with regard to medium-sized municipalities that are represented only by one or two assessments. These internal differences may appear in the study's results as erroneous variance figures that hide the actual links, or at least reduce them.

Because we are dealing with result assessments in the most general sense of the word 'assessment,' all assessments will be treated as equivalent. Each assessment has been standardised on a school-by-school basis; in other words the schools' averages have been converted into z scores, whose average is zero and standard deviation one. In this way we have calculated municipality-specific averages from assessments whose results have been converted into commensurable figures. As we can see from Tables 2 and 3, each assessment has shown regional differences in a different way. The results in English and Swedish seem to be the most notably sensitive subjects to regional differences. The use of averages from the result assessments may not show all the biggest differences, but it increases the thoroughness of the study in two ways: we are able to include more municipalities, and on the other hand we receive a better overview of the municipalities whose schools have taken part in more than one assessment. Thoroughness is considered to be a more important criterion than increasing the sizes of differences or links.

On the basis of their average standard scores, municipalities have been divided into three groups: the lower quartile, the middle 50 per cent (the interquartile group) and the upper quartile. We will examine how the three groups differ from each other with regard to regional background variables. Statistically important criteria are:

1) that the grouping be relevant so that it reflects differences in students' results large enough to consider crucial in terms of developing the educational system
2) that the differences between the demographic factors of the upper and lower quartiles be statistically almost significant when their degrees of freedom are calculated on the basis of municipalities rather than students

3) that the differences be logical; in other words the ranking order must remain in place so that the interquartile group's results will be placed in between the two extremes'
4) that the demographic differences be given a suitable interpretation

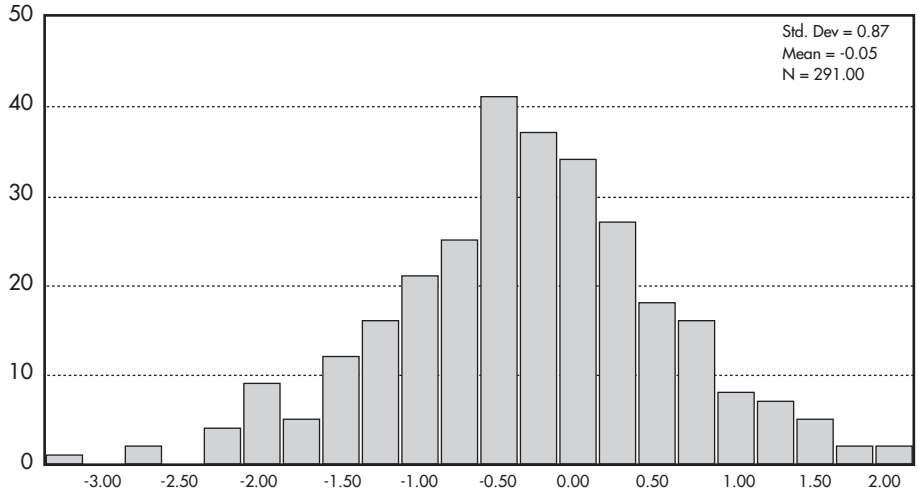
An expected error level of five per cent is considered sufficient, as the number of municipalities is not particularly large, as erroneous variance figures will doubtless appear and as the tests are non-directional.

According to requirement 2 in the above list, it is not a prerequisite that the municipalities in the middle group differ significantly from the other groups, as we have no advance information suggesting that possible connections should be linear. This is why only requirement 3 is logical in that according to it results should be interpreted in a suitably ranked way.

4 RESULTS

The first statistical requirement in the abovementioned list is related to grouping according to students' results and to the relevance of this grouping.

The averages of the standard scores do not quite behave in a standardised manner. The mean of the distribution is -0.05 and its standard deviation is 0.87. The reduction of the standard deviation value can be considered to be a kind of regression towards the mean when there is more than one assessment, but the distribution is still possible to sort. In addition the distribution is slightly unbalanced between left and right and has slightly more excess than normal, but even in those aspects it does not differ in a statistically significant manner from a normal distribution. The two extreme negative values do not matter in material grouped by quartile.



Figur 1. Distribution of the municipalities' standard scores

Table 4. Average performance percentages for municipalities in the different assessments

Municipality group	Maths 98	Sci.	Fin. 99	Eng.	Maths 00	Fin. 01	Swe.
Lower quartile	47%	42%	60%	55%	43%	56%	43%
Interquartile group (middle 50%)	53%	44%	64%	63%	51%	61%	51%
Upper quartile	58%	49%	69%	68%	57%	66%	59%
Total	53%	45%	64%	62%	51%	61%	51%
Difference between upper and lower quartiles	11%	8%	9%	14%	14%	11%	16%

The share point of the lower quartile and the interquartile group is -0.555, and the share point of the middle group and the upper quartile is 0.51. The means and numbers are -1.17 (number= 73), -0.03 (number= 143) and 1.04 (number= 72).

To evaluate whether the division of the municipalities into the three groups also reflects real differences in students' results, we have created a table which describes the performance percentages for each test.

An average difference of 12% can be considered sufficiently significant, especially in the sense that differences of this magnitude have not been found in any of the assessment reports produced by the National Board of Education. Table 4 shows us also that in terms of this index, not all subjects are in quite the same position.

Educational background

In examining the links between education and demographic background factors, it is natural to begin by looking at the educational structure of the data. The population of the upper quartile municipalities is on average more highly educated than that of the lower quartile. As we mentioned earlier, a similar connection has previously been proved at an individual level.

Out of the indicators of the population's level of education, the figures which, on the basis of the education level measurements of Statistics Finland, indicate the level of education of those aged over 15 are very significantly parallel

Table 5. Relationship between the indicators of the population's educational level and the average results of the municipalities

Average performance level – 3 groups	Indicator of level of education of those aged over 15		Portion of qualified individuals		Portion of those who have achieved a mid-level qualification		Portion of those who have achieved a high-level qualification	
	average	z score	average	z score	average	z score	average	z score
Lower quartile	205	-0.259	54	-0.176	39	0.158	16	-0.315
Interquartile group	221	0.146	57	0.179	39	0.152	18	0.127
Upper quartile	235	0.485	58	0.339	38	-0.140	20	0.504
Significance	***		**		n.s.		***	

* p< 0.05 ** p<0.01 *** p<0.001 n.s. p≥0.05

with expectations. The figures that indicate the portion of the population ager over 20 that has achieved a high-level qualification are also very significantly parallel with expectations, whereas the figures regarding the portion of the population aged over 20 that has achieved a qualification of some sort are significantly parallel. The figures related to portion of the population that has achieved a mid-level qualification, however, are not related to the municipalities' students' results. This cannot be considered a particularly unexpected result. The clearest link is between the municipalities' students' results and the portion of adults with a high-level qualification.

The regions' professional and means-related structure and the population's level of income

Within the professional structure of a region, the portion of people employed in managerial positions is the sector that stands out within the professional population. The sectors of entrepreneurs, ordinary and manual workers, students and pensioners have no statistical relation to students' results. On the other hand the proportion of people who do not belong to any of the abovementioned sectors have an inverse relationship with average results.

The population of the municipalities in the upper quartile is more likely to be employed in trade and especially in the worlds of business or finance. The part of the population that is employed in other professions had no effect on our results. In the classification made by Statistics Finland, whose data we have used, 'other professions' include manufacture, industry, construction, transport and other services. On the other hand the relationship between the region's results and the portion of the population that is employed in other unclassified and unknown professions is again inverse.

Table 6. Relationship between the region's professional structure and the average learning-based results of a municipality

	Portion of people in senior management		Portion of people in lower or middle management		Portion of other professions	
Lower quartile	6.6	-0.227	12.6	-0.160	12.9	0.457
Interquartile group	7.8	0.119	13.7	0.129	12	0.130
Upper quartile	9.1	0.540	14.4	0.315	11.2	-0.143
Significance	***		**		***	

Table 7. Relationship between the municipalities' professional structures and their average learning-based results

	Proportion of those employed in trade		Proportion of those employed in business or finance		Proportion of those with unknown professions	
	average	**z score**	**average**	**z score**	**average**	**z score**
Lower quartile	10.8	-0.070	6.6	-0.214	2.9	0.188
Interquartile group	11.4	0.142	7.6	0.146	2.5	-0.137
Upper quartile	11.9	0.288	8.6	0.526	2.4	-0.198
Significance		*		***		*

As indicators of a municipality population's level of income we can use either the average tax rate of the population or the average incomes of households. Out of these two types of indicators, average income figures are more accurate at describing general levels of income, because they do not include for instance the portion that goes towards community taxes. The links found between the average household income indicator and the municipalities' average results were slightly more significant than those found between the average tax rate and the average results.

Table 8. Relationship between the municipalities' average learning-based results and their populations' levels of income

	Average tax rate per person		Average income of households	
	average	**z score**	**average**	**z score**
Lower quartile	54123	-0.109	156862	-0.259
Interquartile group	58043	0.160	163056	-0.012
Upper quartile	60255	0.312	172154	0.351
Significance		**		***

Characteristics of the municipalities

The indicators used so far have in one way or another described the welfare of the inhabitants of the municipalities. Another part of the indicators that are available could be considered more directly to measure the actual status of each municipality.

The classification of municipalities according to EU objective programme areas (earlier called EU subsidy areas), and the subclassification of these into 0 subsidy areas, which receive no subsidies, as well as other areas, has a statistical connection with the classification of municipalities according to students' results. Nine per cent of the municipalities that receive no subsidies are located in the lower quartile of the result-based ranking, whereas 44 per cent of them are located in the upper quartile. These figures for the other objective programme areas are around 30% in the lower quartile and just over 20% in the upper quartile. The expected values for each type of objective programme area are, of course, 25% in the lower quartile, 50% in the interquartile group and 25% in the upper quartile. When compared to the other results obtained from this study so far, this result is consistent in as much as the EU objective programme area classification is also based on a series of factors that describe the community's welfare.

Another classification system related to learning-based results is the statistical classification of municipalities, in which municipalities are divided into urban, suburban and rural areas.

Table 9. Relationship between the European Union's objective programme area classification and municipalities' average learning-based results
($p<0.01$)

		EU objective programme area			
		0	**1**	**2**	**4**
Lower quartile	n	4	32	20	17
	%	9.3	31.7	20.2	35.4
Interquartile group	n	20	49	57	20
	%	46.5	48.5	57.6	41.7
Upper quartile	n	9	20	22	11
	%	44.2	19.8	22.2	22.9
Total	n	43	101	99	48
	%	100	100	100	100

Table 10. Relationship between the statistical classification of municipalities and municipalities' average learning-based results
(p<0.05)

	Statistical classification of municipalities					
	Urban municipalities		Suburban municipalities		Rural municipalities	
	n	%	n	%	n	%
Lower quartile	5	8.3	16	29.6	52	29.4
Interquartile group	37	61.7	25	46.3	84	47.5
Upper quartile	18	30	13	24.1	41	23.2
Total	60	100	54	100	177	100

In this case the main difference can be found between urban municipalities and the other municipalities. The difference would be even more significant if the primary unit of the study were the school and not the municipality, because the very successful schools of the big cities would end up more clearly in the upper quartile than they do in the comparison of municipalities.

In a comparison of larger regions, the district of Etelä-Suomi (southern Finland) appears to do better than the other districts, though the results of the examination of districts do not as a whole show statistically significant differences.

Even though most indicators that are related to municipal economy (debts per inhabitant, cash reserves and relative debt) produce results in the expected direction, only annual balances and tax percentages fulfil, when used as variables, the statistical criteria selected at the beginning. High levels of annual balances and lower-than-average tax percentages and dependency ratios are typical characteristics of the municipalities that have done best in the result assessments.

Table 11. Statistically significant municipal economy-related indicators

	Annual balance FIM/household		Tax rate (%)		Dependency ratio	
	average	z score	average	z score	average	z score
Lower quartile	674	-0.089	18.5	0.479	1.9	0.426
Interquartile group	734	-0.023	18.1	-0.033	1.8	0.024
Upper quartile	1020	0.290	18.0	-0.227	1.7	-0.183
Significance	*		***		***	

Table 12. Statistically significant indicators of population structures

	Density of population inh./km²		Net migration of population %		Net mortality rate		Net immigration Inh.		Densely populated areas %	
	av.	z	av.	z	av.	z	av.	z	av.	z
Lower quartile	20.3	-0.189	-1.04	-0.383	-5.2	-0.161	-53.8	-0.219	56.3	-0.043
Interquartile group	69	0.054	-0.54	-0.009	30.6	0.085	0.3	-0.021	64.3	0.304
Upper quartile	158.8	0.503	-0.22	0.237	61.1	0.294	124.7	0.433	66.1	0.383
Significance		**		***		*		*		**

The size of a municipality's population is by itself not enough to describe statistically significant differences; however we can see that the municipalities in the upper quartile are clearly densely populated areas and receivers of net immigration.

Indicators of regional hardship

The figures for indicators of regional hardship have been taken from the database put together by Ville Viljanen. Viljanen's study deals especially with developments in the 1990s, and he has collected information for the years 1990 and 1998. For this study we have used the data from 1998, and to the table below we have added information regarding the proportion of households in which people live with insufficient space, gathered from the database created by Statistics Finland.

Table 13. Indicators of regional hardship (%)

	Unemployment		Receivers of income support		Those with small incomes		Those living with insufficient space	
	average	z score	average	z score	average	z score	average	z score
Lower quartile	18.2	0.463	10.2	0.512	29.6	0.287	7.6	0.423
Interquartile group	16.5	0.170	9.1	0.188	27.7	-0.074	6.6	-0.049
Upper quartile	14.7	-0.135	8.6	0.028	26.9	-0.231	6.3	-0.189
Significance		***		**		**		***

Out of Viljanen's hardship indicators such variables as psychological or physical crime, alcohol-related healthcare visits or drug-related healthcare visits had no significant effects on students' results in our analysis.

Viljanen has also awarded each municipality a summary score that defines its level of welfare. This score is based on the following variables: net emigration of inhabitants with high-level qualifications in 1998, net emigration between 1996 and 1998 (three-year average), proportion of the population within the optimum working age bracket in 1998, proportion of the population working in the information sector in 1997, proportion of the population that was highly educated in 1998, and changes within high-technology jobs between 1993 and 1997.

For the sake of comparison we formulated, on the basis of results found earlier in this study, a first non-rotating primary component. The variables from which we formed this component are: number of educated inhabitants aged over 15, proportion of population in senior management, proportion of population employed in business or finance, average income of households, net population migration percentage, level of unemployment, and tax rate of the municipality. Let this component be PC1. The two mentioned indices have quite a strong correlation (r = 0.79). Both relate to the classification of municipalities, the second index somewhat more strongly.

Table 14. Two indices: summary score from Ville Viljanen's study and the first non-rotating primary component

	Summary score		PC1
	av.	**z**	**av./z**
Lower quartile	-1.384	-0.265	-0.390
Interquartile group	0.548	0.116	0.065
Upper quartile	2.208	0.443	0.494
	***		***

Summary so far

On the basis of what we have described above we receive a homogeneous picture: good learning-based results are linked to other factors that define regional welfare. Such factors include the population's level of education, and professional and means-based structures. Within variables related to professional and means-based structures, it is worth noting that the issues that were significant were for instance the numbers of people employed in management or working in business or finance. On the other hand no professions or means of subsistence were found which would have had a statistically significant negative effect on students' results. The only statistically significant variables related to bad results were those that had to do with the unclassified portion of the population.

Also high income levels and low rates of communal tax appear to be related to good learning-based results. Good results are on average achieved more often in wealthy, densely populated areas that are net receivers of immigration, than

in other areas. These areas have below-average unemployment levels, and less people live in too-small spaces, have small incomes or need income support.

The mentioned variables are not independent of each other, but it would be impossible to define their mutual relationships in terms of their effect on our study. For instance the adult education level indicators form between then a network of variables in strong correlation. So do the professional and employment indicators, the figures that indicate the population's average income, the figures that measure the municipality's economy and the indicators of regional hardship. These clusters of mutually correlating variables are, furthermore, linked to each other and together form a set of variables that can be seen as describing some of the crucial issues in financial welfare. Education is also a kind of culturally and economically relevant cumulatory asset. It appears that it is regionally focused, just as the other factors that affect welfare. Even though it would be easy to assume that learning-based results probably have no effect on the municipal economy or its net migration levels, the causal relationships between demographic factors and average learning-based results cannot be described so easily. This is why it appears that we would be as justified in considering school results to be just one indicator amongst many of the accumulation of welfare in a region, as we would be in trying to work out the causal relationships between all the variables.

5 METHODICAL FINAL DISCUSSION

Above we have used a very rough approach that causes information to be lost. People acquainted with statistical methods may well ask why a continuous variable such as school results has been subdivided into classes. Why is a continuous variable not treated as continuous and calculated for instance with correlation coefficients?

We wanted to demonstrate the effect that demographic factors have in absolute figures. At times it is quite big, and at other times the average differences seem small. Small differences are mainly due to the fact that deviation is generally small amongst Finnish municipalities. This is why as well as absolute figures we have given z-score averages. On the other hand we must remember that our initial approach was still quite unrefined. The assessment hasn't been regionally thorough enough yet to be able to avoid the kind of erroneous variance that blurs the bigger picture.

The correlation coefficients describe the phenomenon thus:

Table 15. Correlations between certain regional variables, the average learning-based results of all municipalities and the average learning-based results of municipalities that participated in at least three assessments

	All municipalities that took part in the assessments			Municipalities that took part in at least three assessments		
	r	p	n	r	p	n
Summary scores	0.22	0.000	291	0.40	0.001	69
Primary component scores	0.30	0.000	291	0.41	0.000	69
Indicator of education level of population aged over 15	0.26	0.000	291	0.41	0.000	69
Proportion of population with high-level qualification	0.28	0.000	291	0.43	0.000	69
Proportion of population in senior management	0.26	0.000	291	0.36	0.003	69
Unemployment level %	-0.23	0.000	291	-0.36	0.002	69

If a municipality has taken part in more than one assessment, the correlation coefficients rise significantly. The methodical usefulness of the assessment is cumulatory. Accumulation occurs in two ways; the more important of these is the relevance of the demographic data. Municipality-specific data do not necessarily give a correct or accurate picture of the characteristics of the area from which the students are actually from. If the municipality has participated in the assessments more than once it is also more likely that more than one of the municipality's schools be represented in the data. A second, possibly less relevant issue may be related to the fact that random error-causing factors may occur in the assessment, which falsify results in one direction or another.

When evaluating these kinds of models we must take into account three possible sources of errors:

1) Demographic data are not completely reliable. By saying this we are not questioning the validity of the data offered by Statistics Finland or by Ville Viljanen, but how well municipality-specific data can really reflect students' real backgrounds. Errors probably occur in both directions, so they can be considered to be random. They do, however, weaken the explanatory proportions of the model.
2) Related to the tests given to students are random factors, which affect schools' average performances. Such factors could be differences in emphasis in schools' teaching plans or how and when the syllabus given in the teaching plan has been taught. Although the reliability indices of these kinds of tests are usually high, they are not perfect. It is more a part of the assessment task's character to cover the assessed area in a variety of ways, than it is to achieve as high a reliability index as possible.
3) Out of the factors that are not taken into account in the model, the most important one is the kind of variation caused by schools and teaching that we did not even try to take into account in this model. The view according to which demographic factors are the only ones affecting the average performance of a school is overly deterministic, and, as such, grim.

Therefore we can assume that differences between schools depend on background factors that at least partly can be described in demographic terms, *as well as* on the kinds of variations between schools that these background factors do not explain. If out of the three sources of error given above, the first two can be and are minimised (for instance with the help of repeated evaluations), it should be possible to get more information regarding the variations caused by schools and teaching, and the way they should be interpreted.

Acknowledging the existence of the abovementioned sources of error, does not mean, however, that we can question the fact that demographic factors *are* connected to the average results achieved in the area. Erroneous variance will blur the connections that exist rather than emphasise them. The fact that the existence of the connections can be proved regardless of the inaccuracies caused by the data indicates that if more accurate data were available, the connections could be demonstrated even more clearly.

Up to now the extensive teacher and headmaster surveys created by the National Board of Education have not been able to find factors that could explain a significant amount of the differences that exist between schools. One reason for this may be that the variations caused by demographic factors, which bury beneath themselves the effects of the schools' and school officials' operations, must first be eliminated from the differences between schools. Another possible reason is that despite efforts in that direction, the correct questions have still not been asked. A third reason may be that it is actually very difficult to clarify with the aid of questionnaire-based surveys and quantitative analyses what aspects of running a school lead to the achievement of exceptionally good results. Even though it is impossible to carry out a completely thorough qualitative analysis, the situation becomes a whole lot clearer if with quantitative models we can identify the schools which we should most probably be concentrating on when using research and assessment methods that are more suitable for the evaluation of smaller samples.

A correlation coefficient of 0.40 between the education level indicator formulated by Statistics Finland and the average learning-based results of municipalities (Table 16), is perhaps not enough to create an ideal model upon. If we create a regression model whose predictors are the education level indicator for the population aged over 15 (from the Kuntafakta database), the average incomes of households, the tax rate per inhabitant, the density of population, the level of unemployment and the proportion of inhabitants that live in too-small spaces, when a municipality has taken part in at least three assessments we achieve a correlation coefficient of 0.60 between expected and empirical values.

The City of Helsinki has had assessments done of its schools several times now. Thus a uniquely thorough database (by Finnish standards) has been created in that city. Helsinki has also carried out a demographic mapping of the areas students actually come from (rather than where they attend school) in order to work out figures related to so-called positive discrimination. When the learning-based results for various different subjects were combined with demographic

data relating to the areas where students come from, the correlation coefficient obtained was of over 0.70, regardless of the fact that some schools take their students from areas other than the numerically calculated ones [Kuusela, 2000].

In this methodical discussion we have wished to point out clearly that the re-evaluation made in this study is still nowhere near ideal for the purpose of solving the specific evaluation problems related to this issue; however, we can still not conclusively dismiss the results obtained by it. The materials used in this study demonstrate a consequential link between learning-based school results and (other) variables that describe the welfare of a region. If the differences between regions are found and recognised, we can use the data resulting from this study as a tool to aid planning and decision-making. At the same time, in the best possible case we will be able to avoid making naïve generalisations in which average results are seen simply as indicators of a school's 'good' or 'bad' quality. Schools operate in a framework established by social and communal factors. The effect of these factors also on the schools' average results must be acknowledged and recognised before we can begin implementing remedial actions. At the same time we can start off a discussion on whether cumulatory social differences can be evened out simply by using pedagogical or educational policy-related methods, and, if so, on what those methods would be. It is also worth thinking about whether we should consider schools increasingly clearly as a part of society, and not as an educational institution separate and independent of the social structure of its immediate environment. The links are twofold: social differences are reflected in school results, and if on the other hand social differences vary by region, then we can also expect results to vary regionally. If we wish to eliminate the regional differences that exist in school results, we also have to think about socio-political methods of preventing the development of regions' becoming different. If we accept the social differentiation of regions, then we must also obviously accept differences in school results as one way in which that development manifests itself, and see them as a kind of reality that has to be taken into account in educational planning.

If on the other hand we can relate assessment results to the kinds of background factors that describe the demographic operating environment of a school, it is possible for us to begin developing a kind of assessment feedback system by which information given to schools or municipalities gives a better idea of the added value contributed by the school. Even the current preliminary model gives us a chance of looking for positive exceptions to the rule. When doing that we need to figure out whether the appearance of positive exceptions is a case of the municipal background factors not describing demographic effects on education in a relevant way after all (an approach to a greater or lesser degree deterministic), or whether there are factors in the school's operating principles that make it possible for it to achieve good results despite bad starting criteria (pedagogical approach). It is also possible that the pedagogical or school management-related solutions that are successful under socially more difficult conditions will be different from those that prove to work well in more comfortable areas. Then assessment will be a step closer to fulfilling the

developmental task that it has been set, and in the best of cases it will also be more effective (since at least from the point of view of the assessor, we have to consider not only the assessment of effectiveness but also the effectiveness of assessment).

BIBLIOGRAPHY

Hautamäki, J., Arinen, P., Bergholm, B., Hautamäki, A., Kupiainen, S., Kuusela, J., Lehto, J., Niemivirta, M. & Scheinin P. Oppimaan oppiminen ala-asteilla. Opetushallitus. Oppimistulosten arviointi 3/1999. Helsinki.

Hautamäki, J., Arinen, P., Hautamäki, A., Ikonen-Varila, M., Kupiainen, S., Lindblom, B., Niemivirta, M., Rantanen, P., Ruuth, M. & Scheinin, P. Oppimaan oppiminen yläasteilla. Opetushallitus. Oppimistulosten arviointi 7/2000. Helsinki.

Korhonen, H. 1999. Peruskoulun matematiikan oppimistulosten kansallinen arviointi 1998. Opetushallitus. Oppimistulosten arviointi 1/1999. Helsinki.

Korhonen, H. 2001. Perusopetuksen päättövaiheen matematiikan oppimistulosten kansallinen arviointi 2000. Opetushallitus. Oppimistulosten arviointi 3/2001. Helsinki.

Lankinen, M. 2001. Positiivinen diskriminaatio – mitä se on?" Helsingin kaupungin tietokeskuksen keskustelualoite 2001:2.

Lappalainen, H-P. 2000. Peruskoulun äidinkielen oppimistulosten kansallinen arviointi 9. vuosiluokalla 1999. Opetushallitus. Oppimistulosten arviointi 1/2000. Helsinki.

Lappalainen, H-P. 2001. Perusopetuksen äidinkielen ja kirjallisuuden oppimistulosten kansallinen arviointi 9. vuosiluokalla 2001. Oppimistulosten arviointi 6/2001. Helsinki.

Nummenmaa, T., Konttinen, R., Kuusinen, J. & Leskinen, E. 1997. Tutkimusaineiston analyysi. WSOY. Porvoo.

Rajakorpi, A. (toim.) 1999. Peruskoulun 9.-luokkalaisten luonnontieteiden oppimistulosten arviointi. Keväällä 1998 pidetyn kokeen tulokset. Opetushallitus. Oppimistulosten arviointi 2/1999. Helsinki.

Tuokko, E. 2000. Peruskoulun 9. vuosiluokan englannin (A1-kieli) oppimistulosten kansallinen arviointi 1999. Opetushallitus. Oppimistulosten arviointi 3/2000. Helsinki.

Tuokko, E. 2002. Perusopetuksen päättövaiheen ruotsin kielen oppimistulosten kansallinen arviointi 2001. Opetushallitus. Oppimistulosten arviointi 3/2002. Helsinki.

Viljanen, V. 2001. Huono-osaisuuden alueellinen kehitys 1990–luvulla. Suomen kuntaliitto 2001. Acta nro 139. Helsinki.

Välijärvi, J., Linnankylä P., Kupari, P., Reinikainen, P., Malin, A. & Puhakka, E. 2001. Suomen tulevaisuuden osaajat. Koulutuksen tutkimuslaitos, Jyväskylän yliopisto, Opetushallitus & OECD PISA.

Appendix.
REGARDING THE SYNTHESIS OF THE ASSESSMENTS OF FINNISH SPEAKING SCHOOLS IN HELSINKI[1]

The City of Helsinki Education Department and the Finnish National Board of Education have worked together on assessments for several years. In practice this has meant that Helsinki has commissioned all assessments that are carried out on a national basis to be made of all of its schools. The Helsinki Education Department commissioned a report from the National Board of Education in 2001, which was to bring together the results of certain vital assessments and also to relate the average learning results of schools to indicators of so-called positive discrimination. This document is based on that report [Kuusela, 2000].

The indicators of positive discrimination are calculated by catchment area, and they describe a set of central demographic factors which are taken to define a school's operating environment. The indicators have been used as bases for allocating extra resources to schools operating under more demanding circumstances.

The below variables are included in the calculation model developed by specialist Markku Lankinen. Each variable is described in detail in the City of Helsinki Urban Facts Centre's discussion project (2001:2) "Positiivinen diskriminaatio – mitä se on?" ["Positive Discrimination: What is it?" Lankinen, 2001].

- The proportion of families with children that consists of single parent households
- The proportion of households that consists of rented accommodation
- The proportion of those aged over 15 with a low level of education
- The proportion of the city's inhabitants who live in rented accommodation
- The general level of unemployment in the area
- The proportion of those receiving state benefits
- The average income of families with children.

The index is calculated by first making the seven indicators of the area's characteristics comparable. This is done by making each fit to a scale whose average level is zero, and where the common unit of measurement is the standard deviation. This results in 'z points.' The positive discrimination index is the average of these z points. Thus each of the seven variables is equivalent in the structure of the index.

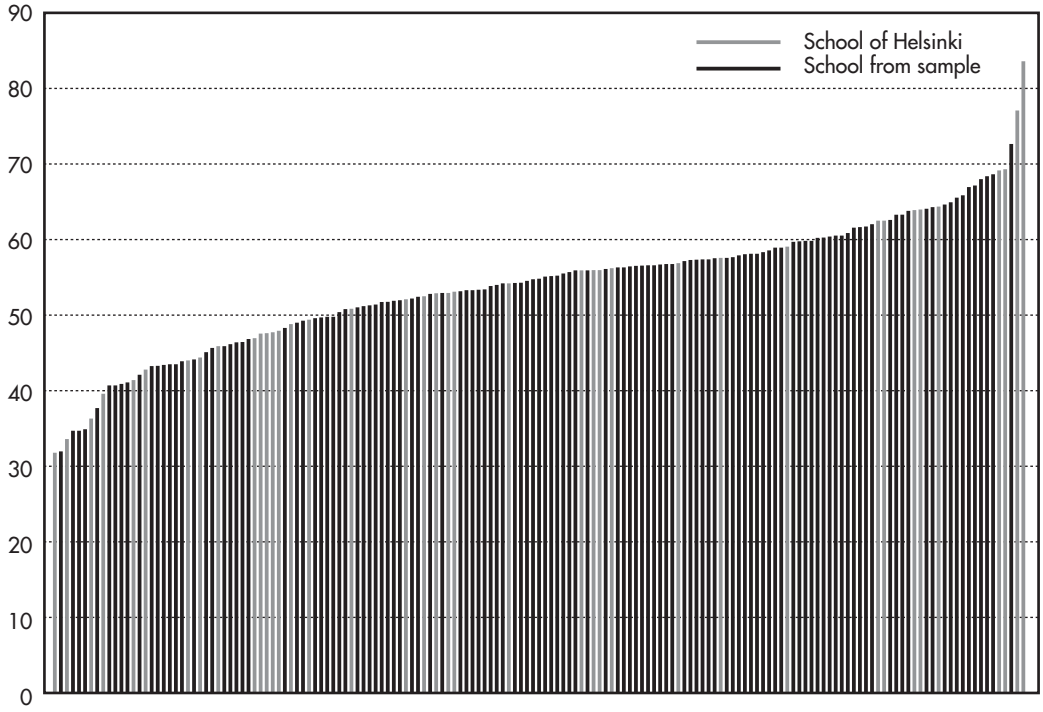
The data on Helsinki are unique for three reasons: they are all-encompassing, they contain repeated assessments, and the results of the schools in them can be related to demographic data based on catchment areas. One thing that causes the data to be inaccurate is still the fact that not all students attend their local schools. The choice of school is a pertinent factor especially within schools that specialise in a certain subject; sometimes the majority of these schools'

1 Published with permission from the City of Helsinki Education Department

pupils come from outside the area considered to be the school's catchment area. Despite this, however, the data on Helsinki are significantly more accurate than the municipality-based national data. The data on Helsinki support two conclusions: on the one hand that there is a strong link between schools' average learning results and the demographic variables affecting pupils' living areas; on the other, that this link becomes increasingly obvious the more accurate the data that is available.

In terms of their average learning results, the upper comprehensive schools in Helsinki are at a similar level to the schools in the national sample group. In general the level of learning results in Helsinki is slightly higher than the average of the national sample. However, the differences between schools in Helsinki are entirely comparable to those between schools in the national sample (Appendix Figure 1).

The report on Helsinki indicates firstly that the schools' average learning results are approximately at the same level regardless of subject or year of assessment. This is essential in terms of the present study. For instance the correlation coefficient of 0.91 shown in Appendix Table 1 between the average results in mathematics and first-language assessments is based on assessments carried out in different years and on different students. This indicates that it is a question more of the average performances of the schools than of results that vary on the basis of subjects or years.



Appendix Figure 1:
Ranking of schools in Helsinki compared to the ranking of schools from the sample group (left blank), sorted by the average performance in mathematics.

Appendix Table 1:
Correlations between assessment results, based on schools' average results.

		mathematics	first language	B-level Swedish
First language	correlation coefficient	0.79		
	significance	0.000		
	N	36		
B-level Swedish	correlation coefficient	0.91	0.91	
	significance	0.000	0.000	
	N	31	30	
A-level Swedish	correlation coefficient	0.59	0.65	0.74
	significance	0.001	0.000	0.000
	N	26	25	20

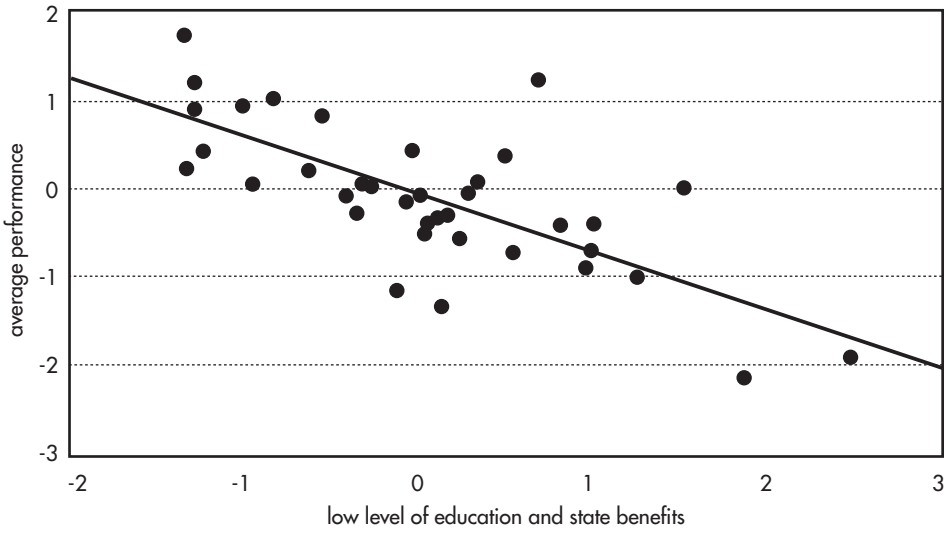
The schools' general results are in strong correlation with the indicators of positive discrimination in the catchment areas. The best predictor for the average results of sixth-graders[2] in mathematics assessments is the positive discrimination index calculated in the way described above. Its correlation to the schools' average learning results is -0.69. If we take four assessments of ninth-graders[3] together, their results are predicted best using a simplified index composed only of the proportion of people with low levels of education and of the proportion of those receiving state benefits in the area.

Appendix Table 2:
The correlation of schools' average learning results with a weighted average index calculated on the basis of the proportion of people with low levels of education and that of those receiving state benefits.

	mathematics	first language	B- level Swedish	A-level Swedish	Average performance from z points
correlation coefficient	-0.63	-0.73	-0.74	-0.57	-0.74
significance	0.000	0.000	0.000	0.002	0.000
number of schools	35	34	29	26	35

When we standardise the four assessments of ninth-graders, and calculate an average to describe a school's average performance level, the correlation of the simplified index with the average learning results is -0.74. The link is strong especially when we take into account the fact that area-based data do not describe the backgrounds of all pupils of all schools equally well. Furthermore, the differences between the average performances of the schools are so large both in Helsinki and in the whole country that this is a link that is significant also in terms of educational policy.

2 Pupils aged 12–13
3 Pupils aged 15–16



Appendix Figure 2:
The correlation between basic education ninth-graders' average performances and the catchment area's low level of education and level of state benefits received.

DISCUSSION

The main building-block of equality in Finnish educational policies is the equality between genders, between regions and between different groups of the population. The ideal situation is one in which each child and teenager reaches a level of basic knowledge satisfactory in terms of further education.

The results of the research indicate that differences do exist between the results of schools that offer basic education in Finland. It has been proved that school-specific differences are related to the schools' operating environments and the demographic indicators of families' levels of welfare. Out of socio-demographic variables, the most important are educational and financial welfare [cf. Kuusela]. Within the learning environment, the level of education of the society and as its financial status are related to a family's and student's security. This is reflected in the level of socio-cultural welfare of the home and the environment in which a child is brought up [cf. Jakku-Sihvonen]. As far as we can see, a higher level of education causes an increased level of culture in a family's everyday life, which, when linked to a good financial situation, leads to increased wellbeing in the home. This has a great significance for the security offered by the environment in which a child is brought up, and for a student's capabilities to concentrate and be motivated in his or her studies.

Reports indicate that some systematical differences exist between schools in Finland. In terms of equality, the most alarming observation is that the average school-specific results of students appear to be related to the different living environments of the population. The most important factors that affect this are financial and educational welfare. In a society that promotes life-long learning, the worst situation is one in which lower levels of education and economy are concentrated in certain areas, affecting the living environments of children and the operation of their schools.

The most important findings of these analyses of basic education results are the polarisation that occurs in certain areas, and the danger of early selectiveness. Polarisation is apparent in the country as a whole, as well as in the capital city area and within Helsinki.

The clear link of the parents' level of education with a student's results increases the danger of social selectiveness, because the selection of a student for continuing from basic into further education promotes the inheritance of higher levels of education. This jeopardises the main target of educational equality, which is that each child and teenager should have the opportunity to choose an educational direction in vocational or college-based learning, which suits his or her skills and inclinations. It is detrimental to the result-orientation of the national educational policies and the equality of students if not all operating environments prepare their students equally well for access into further education.

Indications that the level of results of the upper-quartile schools in northern Finland is lower than that of upper-quartile schools in the capital city area are

worrying, because success in further education is dependent on the skills acquired in basic education.

It is also worrying that differences exist between the performances of boys and girls in the lower-quartile schools. Both first and foreign languages appear to be subjects in which the male students of lower-quartile schools have difficulty reaching even an average level. Boys' performances are poor especially in northern Finland.

In terms of the practical management of schools it is important that we receive as great a variety of information as possible on the causes of the school-specific variations. The information presented in this report is largely related to socio-demographic variables and the description of differences in levels of performance between girls and boys. On the basis of these facts it is possible to direct resources towards the kinds of research and improvement projects that aim to promote the ability of all schools to achieve good results.

A report has recently been published on the international comparative PISA study [cf. Välijärvi *et al.*], which indicates that even in Finland the socio-economic status of the family is related to a student's results. The link, however, is not nearly as significant in Finland as in many other countries.

Right now, the situation is ideal for stemming the development of differences, as the problems have been observed before the differences between population groups or geographical areas have managed to become too large. The task is still not simple, however, as it is a question of preventing the inheritance particularly of low levels of education, and the early creation of gender roles.

An examination of upper-quartile schools shows that girls and boys are capable of achieving similarly high results. Therefore attention must be paid in the home and at school to the attitudes and working practices of children of the two sexes.

It is also appropriate and necessary to focus research on schools that achieve exceptionally good results, against what has been forecasted (based on operating environments). If the cause of the exception lies in the school's pedagogic practices and not in the failing of the forecasts in portraying students' backgrounds, these practices must be noted and used as models to be tested systematically in other, similar schools.

Pedagogic reform measures have to be carefully considered. In connection with this it is worth rectifying the generalisation which unfortunately often appears in the media, according to which learning-based results directly reflect the quality of teaching. We were not able, and did not even attempt to assess the quality of teaching, except in an indirect way with relation to the qualitative nature of the results. The assessments of the National Board of Education have not generally been able to justify students' results using variables related to the quality of teaching, as assessments have been carried out collectively, focusing on schools' results. Still, it is undoubtedly true that if a school achieves above-average results when its background factors give cause to expect below-average results, the teaching in the school is exceptionally high in quality.

In order to develop education, we also need intervention procedures that aim to improve the requirements for students' learning. These kinds of methods

have had good results, and it has been possible to use them to level out even large variations between different groups' starting levels. Some of the intervention programmes in use internationally have been translated into Finnish, and some have even been tried in Finland [cf. e.g. Kuusela, 2000].

One possible method, already implemented in Helsinki, for eliminating geographical differences is so-called positive discrimination: directing extra resources towards schools that need them the most. Even though resources do not automatically transform into results, the combination of extra resources and other measures appears to be the right direction to take.

In our opinions, the results we present here show that there is a need now in educational policy for the kinds of development measures which break the cycle of educational inheritance in society and prevent the polarisation of the population in terms of welfare. It is important to direct educational policy towards promoting equality. From existing Finnish legislation we can extract an ideal of educational equality, by which education reaches everyone; this would mean that living-related and educational welfare were available to all, and that each person had the opportunity to secure his or her livelihood by obtaining employment on the basis of good educational qualifications.