Northern Lights on PISA 2003
– a reflection from the Nordic countries

Edited by Jan Mejding and Astrid Roe

Northern Lights on PISA 2003

– a reflection from the Nordic countries

Edited by Jan Mejding and Astrid Roe

Contents

Foreword

The OECD Programme for International Student Assessment (PISA) is a collaborative effort to measure how well students at age 15 – and thus approaching the end of compulsory schooling – are prepared to meet the challenges of today's societies. PISA combines the assessment of reading, mathematical and scientific literacy with an evaluation of the students' home background and attitudes towards school and learning.

All the Nordic countries – Denmark, Finland, Iceland, Norway and Sweden – participated in PISA in 2000 and 2003. In 2001 the members of the national PISA groups within the Nordic countries decided to prepare a report examining the PISA 2000 results from a Nordic perspective. The report *Northern Lights on PISA* was published in May 2003. Now researchers and policy makers in the Nordic countries have decided to cooperate in publishing a second Nordic report, based on the PISA 2003 results.

All the contributions to this report have been peer reviewed. The report has been funded, supported and published by NSS (Nordisk skolesamarbeid), part of The Nordic Council of Ministers. The editorial meetings were hosted by the Secretariat at the Nordic Council of Ministers. We wish especially to thank Ulla-Jill Karlsson at the Secretariat, who has been very helpful in supporting our work and organising the meetings.

The editorial group

Julius Björnsson
Karl-Göran Karlsson
Pekka Kupari
Jan Mejding
Jørgen Balling Rasmussen
Astrid Roe

Chapter 1

Northern Lights on PISA 2003

An Introduction to the Report

Pekka Kupari and Jukka Törnroos

PISA 2003 was the second survey in OECD's Programme for International Student Assessment. The primary focus of PISA 2003 was on mathematical literacy, with less detailed assessments of science and reading. Problem solving, which was not part of the 2000 survey, was assessed as a minor domain in 2003 but will not be included in later studies. PISA 2003 was conducted in 41 countries, including all 30 OECD countries and 11 partner countries. Generally, the quality standards and procedures for both implementation and reporting results were similar to those of PISA 2000. The first international results of PISA 2003 were reported in 2004 (OECD 2004a, b) and each Nordic country has published its own national report (Kjärnsli et al. 2004; Kupari & Välijärvi 2005; Skolverket 2004c; Mejding 2004; Björnsson et al. 2005).

Mathematics as a major domain

PISA 2003 was based on the approach and methodology described in the OECD framework of assessment (OECD 2003). The assessment focused on *mathematical literacy* by devoting over half of the assessment time to this domain. Mathematical literacy is defined in terms of the capacity of the students to see how mathematics can be used in the real world and thus to engage in mathematics to meet their everyday needs. The mathematics assessment was not simply a test of the students' ability to perform mathematical operations or relate facts. Rather, it was an assessment of how well 15-year-old students recognise, formulate and tackle mathematical problems in the context of real life.

PISA 2003 measured student performance in four areas of mathematics: *space and shape, change and relationships, quantity* and *uncertainty*. Student responses were calculated on 84 different mathematical questions related to the students' personal lives, to education, to work or to issues of wider public relevance. There was not a single cut-off point at which students were deemed mathematically

literate; instead six different levels of mathematical proficiency were used to measure the students' capacity to apply their mathematical knowledge and skills.

What was different in PISA 2003?

PISA 2003 involved at least four aspects or features that were different from PISA 2000 (cf. OECD 2004a). First, in PISA 2003 it was possible to report the students' mathematics performance in much greater detail than was the case in 2000. For the first time performance could be presented in proficiency levels. The results show the percentage of students in each country reaching international benchmarks that measure their mastery of problems at different levels of difficulty. In addition, the reporting scales for mathematics in 2003 were different from the reading scales in 2000: In 2003 four subscales relating to the content areas (*space and shape*, *change and relationships*, *quantity* and *uncertainty*) were used to report results. This kind of reporting allows policy makers to see the way different mathematical competencies have been built up in relation to four broad content areas of mathematics.

Second, the assessment of cross-curricular competencies was extended. In PISA 2000 this domain was explored by asking students about their motivation, self-concept and learning strategies. An important advance was made in PISA 2003 by directly assessing the students' problem-solving skills. Although these skills contribute to performance at school, the problem-solving tasks in PISA 2003 were general, rather than being related to specific curriculum areas. The international results for this domain have been published in a separate report (OECD 2004b).

Third, new background information about students and schools was introduced. The questionnaires explored in greater depth the organisation of school and instructional processes. Focusing on mathematics, students were asked new questions about their attitudes to the subject and their educational careers.

Fourth, PISA allows for comparison over time. PISA measures the students' knowledge and skills in reading, mathematical and scientific literacy according to a basic survey design. This allows countries to see – over time – the effects of policy changes and developmental endeavours on educational outcomes. In reviewing the PISA 2003 results, this possibility must be approached with caution since two sets of results do not demonstrate a trend and since education systems develop relatively slowly.

A brief overview of the PISA 2003 results

Achievement in Nordic and other countries

In the following pages we will present a short overview of the main PISA 2003 results. The overall achievement results in each of the three domains are presented in figures 1-3. In PISA 2003, the performance scale was constructed so that the

average student score for all OECD countries was 500 points and the standard deviation was 100 points (i.e. about two-thirds of students scored between 400 and 600 points). For assessment of reading literacy, the PISA 2003 and 2000 scales were equalised and therefore the mean for all OECD countries was 494 in 2003.

In figures 1-3, the countries are ranked according to their mean scores. The spread of the distribution of scale scores is presented for each country as a standard

Figure 1 *Mean scores and distributions of mathematical literacy*

Country	Mathematics	Mean	S.E.	S.D.
Hong Kong-China		550	(4,5)	100
Finland		544	(1,9)	84
Korea		542	(3,2)	92
Netherlands		538	(3,1)	93
Liechtenstein *		536	(4,1)	99
Japan		534	(4,0)	101
Canada		532	(1,8)	87
Belgium		529	(2,3)	110
Macao-China		527	(2,9)	87
Switzerland		527	(3,4)	98
Australia		524	(2,1)	95
New Zealand		523	(2,3)	98
Czech Republic		516	(3,5)	96
Iceland		515	(1,4)	90
Denmark		514	(2,7)	91
France		511	(2,5)	92
Sweden		509	(2,6)	95
Austria		506	(3,3)	93
Germany		503	(3,3)	103
Ireland		503	(2,4)	85
Slovak Republic		498	(3,3)	93
Norway		495	(2,4)	92
Luxembourgh		493	(1,0)	92
Poland		490	(2,5)	90
Hungary		490	(2,8)	94
Spain		485	(2,4)	88
Latvia *		483	(3,7)	88
United States		483	(2,9)	95
Russian Federation *		468	(4,2)	92
Portugal		466	(3,4)	88
Italy		466	(3,1)	96
Greece		445	(3,9)	94
Serbia		437	(3,8)	85
Turkey		423	(6,7)	105
Uruguay		422	(3,3)	100
Thailand		417	(3,0)	82
Mexico		385	(3,6)	85
Indonesia		360	(3,9)	81
Tunisia		359	(2,5)	82
Brazil *		356	(4,8)	100

200   300   400   500   600   700

* Non-OECD Countries
( ) Standard errors appear in parentheses
International mean = 500

Percentiles

5.   25.   75.   95.

Mean and Confidence Interval (±2 S.E.)

deviation (S.D.). In addition, the 5th, 25th, 75th and 95th percentiles are marked in the bar graphs included in the figures. The dark area in the middle of each bar is the 95 % confidence interval around the mean score. These three figures form the basis for discussion and reference throughout this report.

Figure 2 *Mean scores and distributions of scientific literacy*

Country	Science	Mean	S.E.	S.D.
Finland		548	(1,9)	91
Japan		548	(4,1)	109
Hong Kong-China		539	(4,3)	94
Korea		538	(3,5)	101
Liechtenstein		525	(4,3)	103
Australia		525	(2,1)	102
Macao-China		525	(3,0)	88
Netherlands		524	(3,1)	99
Czech Republic		523	(3,4)	101
New Zealand		521	(2,4)	104
Canada		519	(2,0)	99
Switzerland		513	(3,7)	108
France		511	(3,0)	111
Belgium		509	(2,5)	107
Sweden		506	(2,7)	107
Ireland		505	(2,7)	93
Hungary		503	(2,8)	97
Germany		502	(3,6)	111
Poland		498	(2,9)	102
Slovak Republic		495	(3,7)	102
Iceland		495	(1,5)	96
United States		491	(3,1)	102
Austria		491	(3,4)	97
Russian Federation *		489	(4,1)	100
Latvia *		489	(3,9)	93
Spain		487	(2,6)	100
Italy		486	(3,1)	108
Norway		484	(2,9)	104
Luxembourg		483	(1,5)	103
Greece		481	(3,8)	101
Denmark		475	(3,0)	102
Portugal		468	(3,5)	93
Uruguay		438	(2,9)	109
Serbia		436	(3,5)	83
Turkey		434	(5,9)	96
Thailand		429	(2,7)	81
Mexico		405	(3,5)	87
Indonesia		395	(3,2)	68
Brazil *		390	(4,3)	98
Tunisia		385	(2,6)	87

200 300 400 500 600 700

* Non-OECD Countries
( ) Standard errors appear in parentheses
International mean = 500

Percentiles

5. 25. 75. 95.

Mean and Confidence Interval (±2 S.E.)

Figure 3 *Mean scores and distributions of reading literacy*

Country	Mean	S.E.	S.D.
Finland	543	(1,6)	81
Korea	534	(3,1)	83
Canada	528	(1,7)	89
Australia	525	(2,1)	97
Liechtenstein	525	(3,6)	90
New Zealand	522	(2,5)	105
Ireland	515	(2,6)	87
Sweden	514	(2,4)	96
Netherlands	513	(2,9)	85
Hong Kong - China	510	(3,7)	85
Belgium	507	(2,6)	110
Norway	500	(2,8)	102
Switzerland	499	(3,3)	95
Japan	498	(3,9)	106
Macao - China	498	(2,2)	67
Poland	497	(2,9)	96
France	496	(2,7)	97
United States	495	(3,2)	101
Denmark	492	(2,8)	88
Iceland	492	(1,6)	98
Germany	491	(3,4)	109
Austria	491	(3,8)	103
Latvia *	491	(3,7)	90
Czech Republic	489	(3,5)	96
Hungary	482	(2,5)	92
Spain	481	(2,6)	95
Luxembourg	479	(1,5)	100
Portugal	478	(3,7)	93
Italy	476	(3,0)	101
Greece	472	(4,1)	105
Slovak Republic	469	(3,1)	93
Russian Federation *	442	(3,9)	93
Turkey	441	(5,8)	95
Uruguay	434	(3,4)	121
Thailand	420	(2,8)	78
Serbia	412	(3,6)	81
Brazil *	403	(4,6)	111
Mexico	400	(4,1)	95
Indonesia	382	(3,4)	76
Tunisia	375	(2,8)	96

* Non-OECD Countries
( ) Standard errors appear in parentheses
International mean = 494

Percentiles
5. 25. 75. 95.

Mean and Confidence Interval (±2 S.E.)

The overall results presented in figures 1-3 reveal that when the data for PISA 2003 and PISA 2000 were compared, the mean scores for the Nordic countries remained relatively stable. Only a few significant changes could be identified. In scientific literacy, the mean score in Finland was somewhat higher and the mean score in Norway somewhat lower than 3 years earlier. In reading literacy, correspondingly, the mean score in Iceland was somewhat lower in 2003 than in 2000. Otherwise,

the mean achievement in the Nordic countries can be summarised in figure 4 presented below.

The results show that the Finnish students achieved remarkably high scores in all domains and strongly outperformed their Nordic peers. In fact, in each literacy domain the Finnish students attained the highest scores among all the OECD countries. The performance of the other Nordic students was closer to the OECD mean. In Denmark and Iceland performances in mathematics were better than in science or reading. The performance of the Norwegian students was above the OECD mean in reading, but not in mathematics and science. The Swedish performance profile was similar in 2000 and 2003, above the OECD mean level.

The question of whether a common Nordic profile in mathematics performance can be demonstrated will be analysed further in this report.

Figure 4 *Mean scores in the three literacy domains for the Nordic countries*
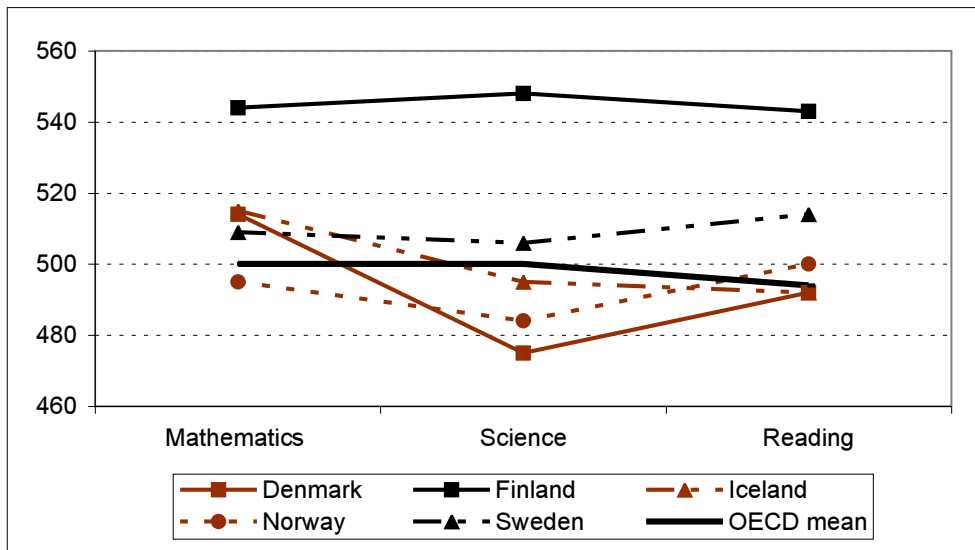
Social equity and literacy achievement

The Nordic countries are well known for their emphasis on social equity and their relative success in this area. Even though the countries are all similar in this regard, differences exist which are important if we want to understand differences in literacy achievement within the Nordic group. Table 1 presents mean mathematics scores for each Nordic country. The mean scores are presented in the second column and the standard deviations in the third column. The fourth column includes the increments in scores associated with an increase in the International Socio-Economic Index (ISEI) of one international standard deviation. The ISEI index is defined as a measure of the socio-economic status of parents, based on

information provided by the students (for details, see OECD 2001, pp. 221). For comparison, the OECD mean values are also presented.

Table 1 *Measures of mathematical literacy achievement: Mean scores, standard deviations and dependence on ISEI index*

	Mean	Standard deviation	Dependence on ISEI
Denmark	514	91	29
Finland	544	84	22
Iceland	515	90	14
Norway	495	92	29
Sweden	509	95	29
OECD mean	500	100	34

The results show that the standard deviations in the Nordic countries were somewhat lower than the OECD total standard deviation[1]. In Finland the standard deviation was one of the smallest among all participating countries.

It is well known that students from less advantaged home backgrounds tend to do less well on average at school than their more advantaged peers. The strength of the relationship between students' socio-economic background and their mathematics performance can be seen from the table. The results show that among the OECD countries the relationship was weakest in Iceland and Finland. In the other Nordic countries the relationship was also weaker than in the OECD countries on average. The results were very similar to the results for reading literacy in PISA 2000.

Table 2 *Gender differences in score points within the three literacy domains. Negative differences are in favour of girls*

	Mathematics	Science	Reading
Denmark	17	17	-25
Finland	7	-6	-44
Iceland	-15	-10	-58
Norway	6	2	-49
Sweden	7	5	-37
OECD mean	11	6	-34

1. The OECD total standard deviation (100) includes the within country standard deviation (94) and the between country standard deviation (6).

Table 2 presents the gender differences in the three content domains for the Nordic countries and for the OECD countries as a whole. Some clear observations can be made on the basis of the results. The gender gap was largest in reading literacy and was in favour of girls just as in PISA 2000 (cf. Lie et al. 2003). In all Nordic countries except Denmark the gap was wider than in the OECD countries on average. Further, in mathematical literacy the overall difference was not large, even though boys outperformed girls in most countries. In the Nordic countries the gender gap in favour of boys was largest in Denmark. The gender gaps in Norway, Sweden and Finland were some of the smallest among all OECD countries. The gender gap in Iceland, on the other hand, was a striking exception, being wide and in favour of girls. In scientific literacy there were no systematic differences between boys and girls. This was an interesting result because boys have generally performed better than girls in science in the past. These issues will be further analysed and discussed in this report.

The scope of the present report

The chapters of this report are based on analyses made by researchers involved in PISA 2003 in the Nordic countries. The chapters have been written with the intention of giving policy makers, researchers and teachers useful insights beyond the simple ranking lists that always receive most attention when the results of PISA are published. This report illuminates many different ways in which analyses of the PISA data can contribute to the educational field. Additionally, many of the chapters address the question of whether or not we can identify a common Nordic profile in the PISA results. The Nordic countries certainly have much in common economically, historically and culturally, but what similarities are revealed through the lens of an international study?

The report can be roughly divided into four thematic parts: In the first part various aspects of the results in mathematics are presented. The chapters by Kupari & Törnroos, Olsen, and Olsen & Grønmo discuss characteristics of the performances in mathematical literacy in the Nordic countries in PISA 2003. The conclusions relating to the existence of a Nordic profile vary depending on the analytical methods used in the studies: Finland is clearly different from the other Nordic countries when levels of performance are studied (Kupari & Törnroos), but the relative strengths and weaknesses of the Nordic countries seem to be similar (Olsen). Moreover, at the item level the performance profile of the Nordic countries is closely connected to whether or not the items represent 'realistic mathematics' (Olsen & Grønmo). Allerup, Lindenskov & Weng study the use of double-digit coding in the marking of students' answers to open-constructed items.

The second part of the report consists of chapters that describe different kinds of background variables and their relation to student achievement. Lie & Kjærnsli

state that when factors related to students' learning strategies, motivation, self-efficacy, classroom and school contexts, and home background are studied, the Nordic countries, particularly Finland, Norway and Sweden, form a distinct group compared with the other countries that participated in PISA. Törnroos, Ingemansson, Pettersson & Kupari look more closely at four affective factors (students' self-concept, interest, motivation, and anxiety in mathematics). According to them, students' attitudes towards mathematics vary between the Nordic countries, but the connections to performance are nonetheless similar. Turmo & Nerheim Hopfenbeck analyse students' learning strategies and their connection with performance. Their analyses show, for example, that students in the Nordic countries do not make use of different kinds of learning strategies as much as their peers in the OECD countries on average. Roe and Hvistendahl examine the results of minority students in the Nordic countries. Unsurprisingly the minority students achieve lower scores than the majority students; however, their results vary considerably between different Nordic countries and areas of literacy (mathematics, reading and scientific), as well as between students who were born in and outside the countries.

In the third part of the report reading and scientific literacy in the Nordic countries is discussed from various points of view. Roe & Taube study the connection between reading and mathematical literacy in PISA 2003. They claim, for example, that the level of understanding and interpretation of verbal expressions needed for the mathematics items differs between items, depending on the correlation between the items and the overall reading score. Linnakylä, Malin & Taube analyse socio-cultural factors related to increased risk of low reading literacy proficiency in Finland and Sweden. According to them, male gender, immigrant status, low socio-economic background, lack of educational and cultural resources at home, and low educational aspirations are the main factors increasing the risk in both of the countries. Leino & Malin study the relationship between use of ICT and reading literacy achievement. Their results show that particularly boys' reading literacy proficiency could benefit from use of ICT. Karlsson, Kjærnsli, Lie & Åström present the only chapter related to scientific literacy in this report. They discuss the changes in students' competencies in science in Norway and Sweden between years 1995 and 2003 and how these changes are related to the educational reforms made recently in these countries. In their analyses the authors also utilise data from the TIMSS studies (Third International Mathematics and Science Study).

The last group of chapters consists of in-depth analyses of the PISA 2003 results at the national level and presentations of national extensions to the international PISA study. In their chapter, Ólafsson, Halldórsson & Björnsson study the exceptional gender difference in favour of girls in mathematics literacy in Iceland. They also take a look at the urban-rural differences in Iceland. They suggest that the gender difference in mathematics may depend on the gender differences also

seen in reading literacy, but this needs to be studied further in the future. Mejding, Reusch & Yung Andersen explore the connections between the Danish school-leaving examination marks in mathematics, Danish, and physics and chemistry on the one hand, and the PISA results in mathematical, reading and scientific literacy on the other hand. The relatively high correlations between these estimates confirm that PISA measures essential skills from the school's point of view. The last two chapters of this group present two national extensions of PISA in Denmark. Egelund and Rangvid present results from PISA Copenhagen, a municipal school development project. The results of the project show, for instance, that the social background and immigrant status of the families living in the school district heavily affect the schools' outcomes. Data from another extension of the original PISA programme, the PISA longitudinal database in Denmark, are presented by Jensen and Andersen. The goal of this project is to illuminate the path from childhood to adulthood and the important role of education in building this path. One of the results so far is that reading skills are an important predictor of future career choices.

Concluding remarks

This report is not just for researchers. Our main goal has been to communicate interesting findings from the PISA 2003 data to policy makers and educators. It is of course no simple matter to draw political or educational implications from the wealth of information produced in the PISA studies. An important aim of the report has been to show that educational achievement in general cannot be understood in terms of simple relationships between single variables. Educational results are multivariate in nature and simplistic conclusions should be treated with suspicion. Drawing policy implications from the wealth of information available is a difficult task for our policy makers, and this report aspires to assist with that task. The PISA cycle of studies gives a general although perhaps narrow picture of the status of our educational systems and provides a small glimpse into the future, or at least some indication of where we are going. The Nordic countries are generally doing well, and some of them are doing exceptionally well, as the Finnish results show. There are similarities between our countries which can help us all understand better how our education systems work and the differences between us are also illuminating in many respects.

   This report has only touched upon some of the very important issues concerning educational achievement and the focus of the report is largely the result of the individual author's interests and expertise. Many important issues have not been covered and many questions have been left unanswered, but this is the nature of every scientific endeavour. The researchers who have collaborated in this work have all learned from the exercise and their general experience has been that this

kind of Nordic collaboration is a unique and fruitful way of gaining extra insights into each individual country's results. International comparative research is thus helping us all to understand our own educational systems better, validating our results and conclusions and helping the future development of education in our countries.

References

Bjornsson, J, Halldorsson, A.M og Olafsson, R.F. Stærdfrædi vid lok grunnskóla: Fyrstu niðurstöður PISA 2003 rannsóknarinnar. Námsmatsstofnun 2004. (Icelandic PISA 2003 report).

Kjærnsli, M., Lie, S., Olsen R. V., Roe, A. og Turmo, A. (2004). Rett spor eller ville veier? Norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003. Oslo: Universitetsforlaget.

Kupari, P. & Välijärvi, J. (eds.) (2005). Osaaminen kestävällä pohjalla. PISA 2003 Suomessa. Jyväskylän yliopisto. Koulutuksen tutkimuslaitos.

Lie, S., Linnakylä, P. & Roe, A. (eds.) (2003). Northern Lights on PISA. Unity and diversity in the Nordic Countries in PISA 2000. Oslo: University of Oslo.

Mejding, Jan (red): "PISA 2003 – Danske unge i en international sammenligning", DPUs forlag, København, 2004

OECD (2001). Knowledge and skills for life. First results from PISA 2000. Paris: OECD.

OECD (2003). The Pisa 2003 assessment framework. Mathematics, reading, science and problem solving knowledge and skills. Paris: OECD.

OECD (2004a). Learning for tomorrow's world. First results from PISA 2003. Paris: OECD.

OECD (2004b). Learning for tomorrow's world – First measures of cross curricular competencies from PISA 2003. Paris: OECD.

Skolverket (2004c). PISA 2003 - Svenska femtonåringars kunskaper och attityder i ett internationellt perspektiv. Rapport 254. Stockholm: Skolverket.

Chapter 2

Characterising Students' Mathematical Literacy Performances in Nordic Countries

Pekka Kupari and Jukka Törnroos

Abstract

Various analyses of PISA 2000 science and reading literacy data indicate that there are some similarities but also differences in students' performances among the Nordic countries. The present article examines 15-year-old students' mathematical literacy performances in five Nordic countries in the PISA 2003 study and explores the similarities and differences detected. The national data for these five countries were analysed from two perspectives: (1) students' levels of performance and performance profiles and (2) gender differences in performance. Relevant OECD averages and the performance profiles of four other countries were used as comparison benchmarks for defining the similarities and differences in Nordic students' mathematical literacy performance.

Nordic abstract

PISA 2000 –aineistosta tehdyt luonnontieteiden ja lukutaidon tarkastelut indikoivat, että oppilaiden suorituksista löytyy yhtäläisyyksiä mutta myös erilaisuutta Pohjoismaiden välillä. Tässä artikkelissa tarkastellaan 15-vuotiaiden nuorten matematiikan suorituksia Pohjoismaissa PISA 2003 –tutkimuksen aineiston pohjalta, ja etsitään suorituksissa ilmeneviä samanlaisuuden ja erilaisuuden piirteitä. Viiden Pohjoismaan kansallisia matematiikan suorituksia analysoitiin kahdesta näkökulmasta: yhtäältä kuvaten suoritustasoa ja oppilaiden suoritusprofiileja ja toisaalta tarkastellen suoritusten sukupuolieroja. Kunkin Pohjoismaan kansallisia matematiikan suorituksia verrattiin OECD-maiden keskitasoon ja osin myös neljän ulkopuolisen vertailumaan vastaaviin suorituksiin. Tulokset osoittivat, että matematiikan suorituksissa on Pohjoismaiden välillä samankaltaisuuksia, mutta myös selkeitä eroja. Suomalaisoppilaiden korkea suoritustaso näyttäisi olevan paljolti seurausta siitä, että heikommin menestyneet oppilaat olivat huomattavan paljon edellä muiden maiden vastaavia oppilaita.

Introduction

In this article we will analyse 15-year-old students' mathematical literacy competence in five Nordic countries on the basis of the PISA 2003 data. Our purpose is to identify features in the performance data that reflect both similarities and differences among these countries. For the purposes of a comparison of mathematics achievement among the Nordic countries, PISA 2003 provides a particularly interesting basis for a number of reasons. Firstly, the analyses of PISA 2000 data on reading and scientific literacy (e.g. Lie & Roe, 2003; Kjärnsli & Lie, 2004) provide evidence of obvious similarities among the Nordic countries. At the same time these analyses reveal that Finland is different in some respects and is atypical of the group. Secondly, not all the Nordic countries took part in earlier international assessments (e.g. TIMSS 1995, 1999) at the same time, so that any direct comparisons of their mathematics achievement were then impossible. Moreover, in PISA 2003 mathematics was the main assessment domain, and it therefore also offers representative and rich data for analysis. Thirdly, PISA assesses mathematics achievement in terms of mathematical literacy. This approach highlights the students' capacity to apply and use learned mathematical skills and knowledge in situations that are as authentic and meaningful for their future as possible (OECD, 2003).

This article tries to answer two questions: What kind of similarities and differences in mathematical competencies are there between the Nordic countries? Is there a common pattern in Nordic students' mathematical literacy performances? In addition, we will offer some interpretations of the differences and similarities detected.

To find answers to the questions we will analyse and compare the national data on mathematical literacy performance among five Nordic countries (Denmark, Finland, Iceland, Norway and Sweden) from two perspectives. On the one hand, we will look at the performance profiles of students at different achievement levels, both across and within content areas. On the other hand, we will compare gender differences in mathematical literacy in the light of students' overall scores and item-specific results. For comparison benchmarks we will use the relevant OECD averages as well as the performance profiles of four other countries (Hong Kong-China, Hungary, Canada and Netherlands). These countries are selected to represent different kinds of school culture in terms of their respective performance levels and educational systems.

PISA and mathematical literacy

The PISA 2003 mathematics assessment was based on three elements or dimensions: *content, processes,* and *situations.* Mathematical content was defined in terms of four broad areas: *quantity, space and shape, change and relationships,* and *uncertainty.* According to the PISA 2003 framework (OECD, 2003, pp. 36-37):

- *Quantity* involves numeric phenomena as well as quantitative relationships and patterns. It relates to the understanding of relative size, the recognition of numerical patterns and the use of numbers to represent quantities and quantifiable attributes of real-world objects.
- *Space and shape* relates to spatial and geometric phenomena and relationships. It requires looking for similarities and differences when analysing the components of shapes and recognising shapes in different dimensions, as well as understanding the properties of objects and their relative positions.
- *Change and relationships* involves manifestations of change as well as functional thinking and dependency among variables. Mathematical relationships are often expressed as equations and inequalities, but relationships of a more general nature are relevant as well. Relationships are given a variety of different representations, including symbolic, algebraic, graphical, tabular and geometrical representations.
- *Uncertainty* involves probabilistic and statistical phenomena and relationships. Collecting data, data analysis and visualisation, probability and inference are important mathematical concepts and activities in this content area.

Mathematical processes were categorised into three clusters: *reproduction, connections,* and *reflection* (OECD, 2003, pp. 42-47):

- The *reproduction* cluster essentially involves reproduction of practised knowledge. The most common competencies are knowledge of facts and of common problem representations, recollection of familiar mathematical objects and properties and performance of routine procedures and standard algorithms.
- The *connections* cluster builds on reproduction to solve problems that are not simply routine, but still involve familiar settings. Problems typically involve greater interpretation demands and require making links between different representations of the situation.
- The *reflection* cluster builds further on the connections cluster. These competencies are required in tasks that demand some insight and reflection on the part of the student, as well as creativity in identifying relevant mathematical concepts or in linking relevant knowledge to create solutions.

An important aspect of mathematical literacy is engagement in mathematics, which means using and doing mathematics in a variety of *situations*. In PISA 2003 there

were four sorts of situations: *personal, educational or occupational, public,* and *scientific*.

There were 84 mathematics items altogether in PISA 2003, of which two-thirds were open ended and the rest multiple-choice format. The items were mostly distributed evenly across the four content areas, except for the process categories, which were less evenly distributed. The largest category was the *connections* cluster (39 items), and the *reflections* cluster was the smallest (19 items).

Results

In PISA 2003 the mathematics performance scale for illustrating and comparing the results was constructed so that the average score of students' performances across all OECD countries was 500, with a standard deviation of 100 points. This meant that about two-thirds of students among OECD countries had scores between 400 and 600 points (OECD, 2004). In addition, seven *proficiency levels* were defined in order to allow for more detailed descriptions, so that the width of one level on the performance scale was always 61 score points. Corresponding scales and proficiency levels were constructed for each content area.

Overall performances

Figure 1 gives an overall picture of mathematical literacy achievement in the Nordic countries compared with the OECD average. As the figure shows, all Nordic countries except Norway performed better than the OECD average. The performance level of Finnish students was the highest of all OECD countries and well above the other Nordic countries, while performance levels in Iceland, Denmark and Sweden were reasonably close to each other. The relative performance levels of the Nordic countries were very much the same for both total scores and results for different content areas. The content area of *uncertainty*, however, formed an exception: in this area the Norwegian students also performed above the OECD average and Finland's advantage over the other Nordic countries was smaller. Furthermore, the standard deviations of the national total scores for all Nordic countries were below the OECD average (100), ranging from 84 (Finland) to 95 (Sweden) score points. The overall results raise one question above all: what explains such remarkable differences in mathematical literacy achievement between the Finnish and other Nordic students. We will investigate this question by analysing the performance profiles in more detail.

Figure 1 *Mathematical literacy performance in the Nordic countries relative to the OECD average*
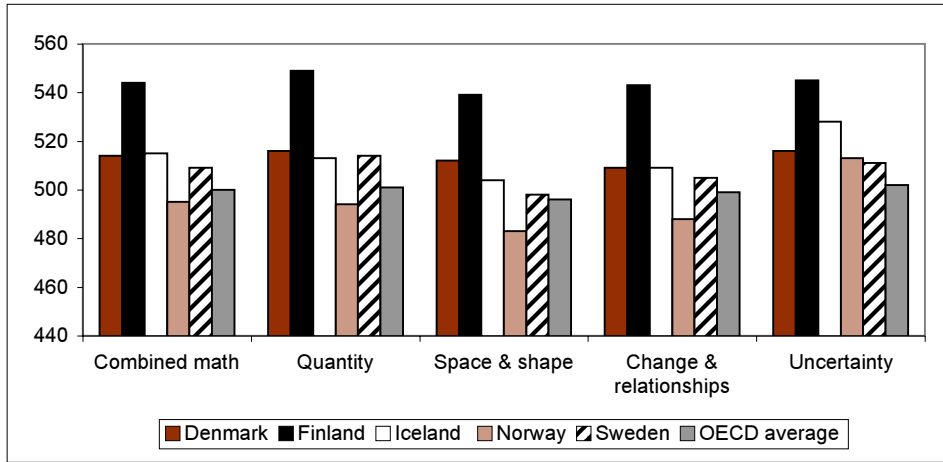
Performance profiles

We will next take a closer look at the mathematical literacy performance of students at different levels. This is based on the percentiles of national performance data (seven percentile points), which are normalised to the corresponding OECD averages (by subtraction) and illustrated by profiles drawn accordingly. Figure 2 presents performance profiles for the Nordic countries and four other countries for comparison (the 0 level represents the OECD average).

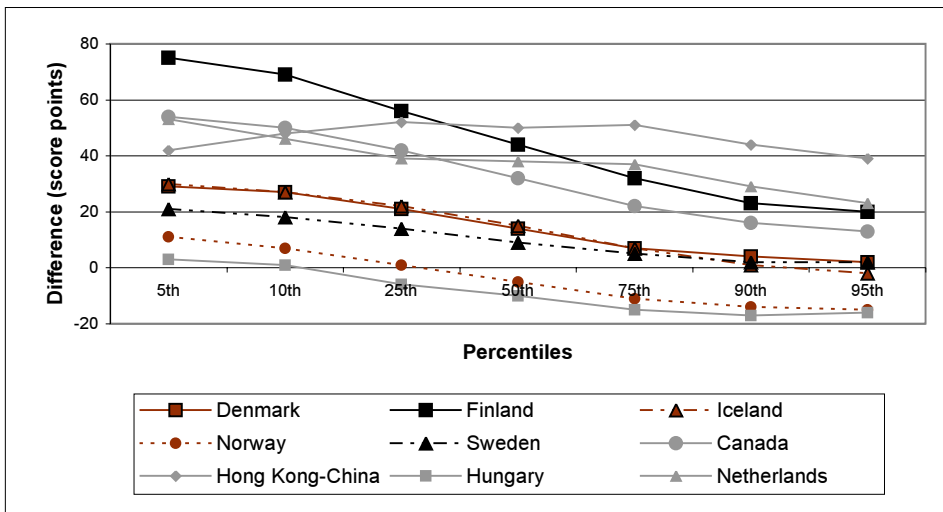**Figure 2** *Profiles of mathematical literacy performance in the Nordic countries*

Figure 2 reveals some very interesting results. The Finnish performance profile shows that the mathematics scores of the lowest achieving students' (5th and 10th percentile) were particularly high and were even clearly above the results of Hong Kong-China, which was on average the best performing country in PISA 2003. These Finnish students achieved about 70 score points above the corresponding OECD average (more than one proficiency level), and also some 40 to 60 points above the scores of the other Nordic countries. In contrast, within the group of highest-achieving students (90th and 95th percentile) the Finnish students' results were only 20 to 30 higher than those of their Nordic peers and remained below the level of Hong Kong-China and the Netherlands. The performance profiles of the other Nordic countries have a slightly gentler contour compared to the Finnish profile. The Danish, Icelandic and Swedish profiles are very much alike, whereas the Norwegian profile is close to that of Hungary in this comparison. When performance profiles were drawn in a similar fashion for specific content areas they showed that some features varied by area. The data in Table 1, where score differences at three percentile points are shown, should help the reader to see the differences between the content areas among the Nordic countries.

Table 1 *Percentile score differences among the Nordic countries (relative to the OECD average) at three percentile points in four content areas of mathematical literacy*

	Percentile score differences (relative to OECD average)											
	Quantity			Space & shape			Change & relations.			Uncertainty		
Country	10th	50th	90th	10th	50th	90th	10th	50th	90th	10th	50th	90th
Denmark	29	15	3	26	16	4	26	10	-3	22	14	3
Finland	75	48	25	67	43	19	66	44	27	63	43	23
Iceland	20	12	4	26	8	-17	26	10	-4	31	26	18
Norway	6	-7	-15	-4	-13	-24	4	-11	-24	12	11	11
Sweden	32	13	-1	17	2	-12	6	6	11	10	9	11

Table 1 reveals that the lowest-achieving Finnish students had high results in all content areas of mathematical literacy. Further, it can be seen that the Danish and Finnish profiles were consistent in all content areas. An interesting feature in the area of *space and shape* was that the profiles for Iceland and Sweden showed a steep decline at the upper end of the distribution (90th percentile), falling below the OECD average level. In addition, the table shows that Norwegian scores were consistently below the OECD average.

When it came to variation within the national profiles, the biggest disparity could be seen in the Swedish profile for the content area of *change and relationships*; the profile was quite flat, starting close to the OECD average at the lower end and finishing at a higher level, somewhat above the Danish and Icelandic profiles, at the upper end of the distribution. In the area of *uncertainty* the performance profiles of

Iceland, Norway and Sweden were even flatter: they ran almost horizontally, at between 10 and 30 score points above the OECD average.

The data for Canada, Hong Kong-China, the Netherlands and Hungary are not shown in Table 1 but it is worth mentioning that in all content areas the Finnish profile followed those of Canada and the Netherlands. Another interesting finding was that the performance profiles of Hong Kong-China across different content areas showed little change. The shape of the profiles stayed almost identical compared to that in Figure 2, although there were slight changes in position, with the profile for *space and shape* at the top and the profile for *change and relationship* at the bottom.

Gender differences

The PISA 2003 mathematics results showed that three Nordic countries, Norway, Sweden and Finland, were among the group of OECD countries where gender differences were smallest. Boys performed slightly better than girls in all countries except Iceland. The difference in mean scores was 6 points in Norway and in Finland and Sweden it was 7 points, which was a statistically significant difference. In Denmark the difference was considerably greater: boys were17 score points ahead of girls. Iceland was unique among the OECD countries, with the girls outperforming boys by 15 score points. This issue is discussed in more detail in chapter 14 (Ólafsson et al.).

These results indicate that there are distinct similarities and differences between the Nordic countries regarding the relative mathematical literacy performances of girls and boys. In the following section we will take a closer look at the nature and structure of these gender differences. Are the differences consistent across the various proficiency levels, content areas and item-specific process clusters?

Table 2 presents the percentages of girls and boys at different proficiency levels of mathematical literacy in each of the Nordic countries and in the OECD countries on average.

The data show that the gender differences vary considerably within the Nordic countries in the PISA 2003 results. The better performances of Danish boys and Icelandic girls are apparent. In Denmark boys were in the majority at the higher proficiency levels (Levels 3 through 6) but in the minority at the three lower levels. In Iceland, in contrast, girls dominated at the higher levels, especially at Level 4, while boys were in the majority at the two lowest levels. The percentages of Icelandic girls at the lowest proficiency levels were well below the OECD averages. In Finland the distribution of the genders across the performance scale was

Table 2 *Percentages of girls and boys at different performance levels of mathematical literacy in the Nordic countries*

Level	Denmark		Finland		Iceland		Norway		Sweden		OECD	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Level 6	4.9	3.3	8.2	5.1	3.7	3.8	3.5	1.9	4.9	3.4	5.1	2.9
Level 5	13.1	10.6	17.7	15.7	11.4	12.2	9.7	7.7	12.4	10.9	11.8	9.5
Level 4	23.5	20.4	25.4	26.9	21.0	25.5	19.1	18.7	19.4	20.2	19.5	18.8
Level 3	26.4	26.0	25.9	29.5	25.3	26.9	23.9	26.5	25.4	25.6	22.9	24.5
Level 2	18.7	22.3	15.4	16.7	20.4	20.1	23.2	24.1	21.3	22.1	20.0	22.1
Level 1	9.6	11.8	5.8	4.9	12.1	8.8	13.3	14.5	11.1	12.3	12.6	13.8
Below level 1	3.8	5.6	1.6	1.4	6.1	2.8	7.3	6.5	5.6	5.6	8.1	8.4

different; boys were in the majority at both ends of the scale, i.e. at the two highest and the two lowest proficiency levels, while girls outnumbered boys at Levels 2 through 4. In Norway too there were more boys than girls at the two highest proficiency levels, but the percentages remained below the OECD average. There were also more boys than girls at the lowest proficiency level, whereas girls were in the majority at Levels 1 through 3. In Sweden the distribution of genders across the proficiency levels was close to the OECD average distribution. Boys were in the majority at the two highest levels, while girls outnumbered boys at Levels 1 through 4.

Figure 3 shows that within the Nordic countries the gender differences in students' mathematical literacy performance also varied considerably by content area.

Denmark and Iceland were at opposite ends of the range of gender differences. The figure shows that where Denmark had the smallest difference favouring boys Iceland had the greatest difference favouring girls, and vice versa. In the area of *quantity* the gender gap in Iceland was about 30 score points (equivalent of half a proficiency level), while the gap for Danish students in the area of *uncertainty* was about 20 points (a third of a proficiency level). In Finland, Norway and Sweden the gender differences were below the OECD average in all content areas. In Finland the gender difference was negligible in the areas of *quantity* and *space and shape,* and for the latter area the difference was notably below the OECD average. In Norway and Sweden the gender differences were almost identical in all content areas and in the area of *change and relationships* the differences were much smaller than in the other Nordic countries.

Figure 3 *Gender differences within different content areas (positive values favouring boys)*
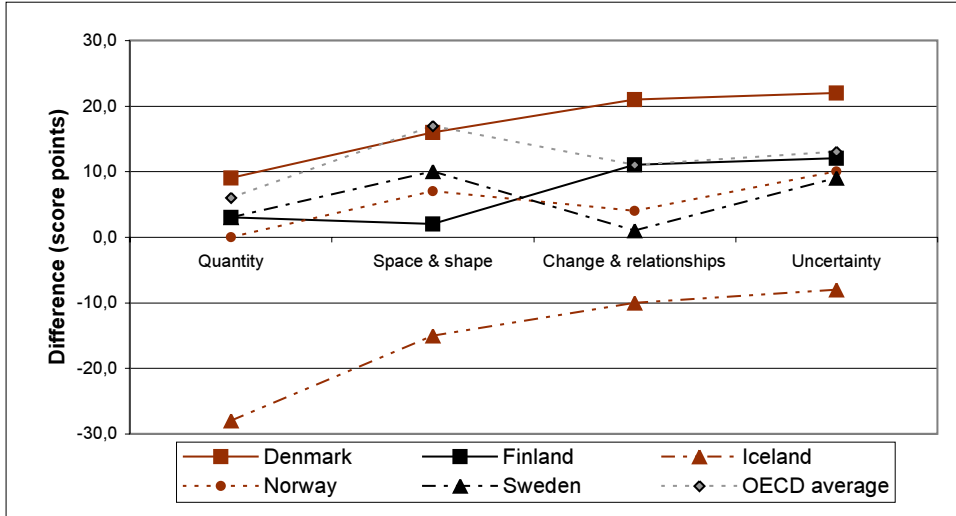
Finally, we will examine boys' and girls' performances with respect to the mathematics tasks in the different process clusters. The PISA 2003 assessment involved three process clusters: a *reproduction* cluster (26 items), a *connections* cluster (39 items) and a *reflection* cluster (19 items). Because no performance scales were constructed for these process clusters, the analysis is based on item-specific response statistics (facility means).

Table 3 *Facility means (percent correct) and gender differences of mathematical literacy items in different process clusters in the Nordic countries*

Country	Reproduction		Connections		Reflection	
	Mean	Gen diff	Mean	Gen diff	Mean	Gen diff
Denmark	72.0	2.8	50.0	3.4	36.5	2.8
Finland	75.7	0.5	56.0	1.7	42.8	1.4
Iceland	69.7	-3.5	50.8	-3.7	36.4	-2.5
Norway	67.0	0.5	44.8	1.9	34.5	1.9
Sweden	68.3	1.7	49.0	2.3	35.7	1.2

These data are understandably well correlated with the findings based on score points. Finnish students had the highest facility means in all process clusters; otherwise the country averages were relatively similar. The fact that the Norwegians performed slightly below the OECD average seems to stem largely from the poor performance of Norwegian students in items belonging to the *connections* cluster. The table reveals clearly how the demand hierarchy between the items in different process clusters worked in each country. The decline in facility means between the *reproduction* and the *connections* clusters was about 20 percentage points and a further 10–15 percentage points between the *reflection* and the *connections* clusters.

The information in Table 3 also seems to indicate that in all Nordic countries except Iceland boys performed slightly better than girls in each process cluster. We should look more carefully at the data, however, because this table does not account for the simultaneous effect of content areas on these facility means. An analysis of the items in different process clusters by content area yielded very interesting findings. Figure 4 displays the results for the *reflection* cluster.

Figure 4 *Gender differences of reflection cluster items in different content areas (positive values in favour of boys)*
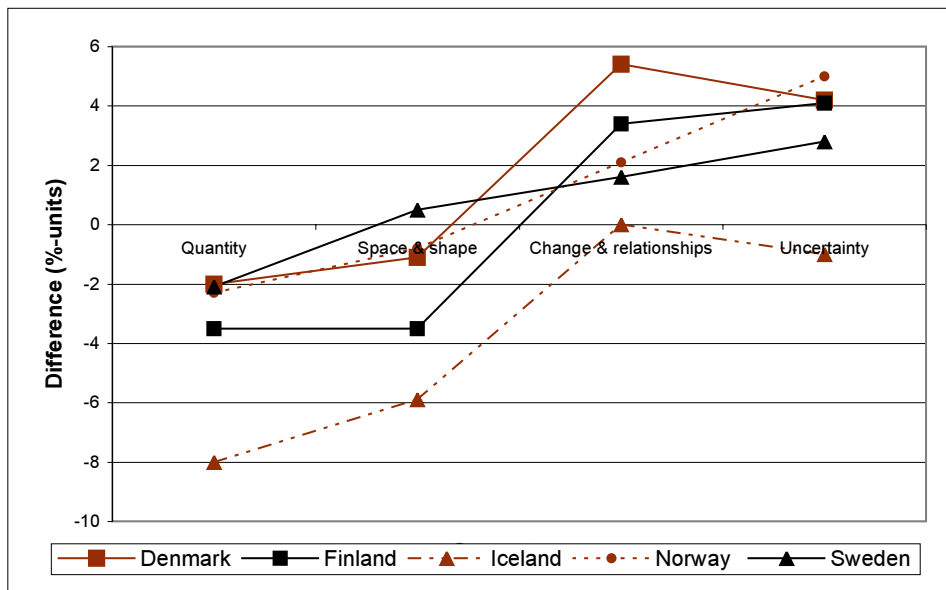
The figure shows that the average gender differences do vary considerably depending on the content area represented in the *reflection* items. In all five Nordic countries girls performed better than boys in the area of *quantity* and also, apart from Sweden, in the *space and shape* items. In contrast, in the areas of *change and relationships* and *uncertainty* boys did better than girls in all Nordic countries except in Iceland. Although in Iceland the gender differences were minimal in these areas. The shape of the profiles in the figure indicates similar trends in all Nordic countries.

Discussion

This article analyses students' mathematical literacy performance in the PISA 2003 study in five Nordic countries, whose education systems and education policies have much in common (e.g. Gorard & Smith, 2004; Husén, 1974). The similarities and differences in the Nordic performances have been examined in relation to OECD averages and also to the data for four other PISA-countries used as benchmarks.

Based on the analyses, there seem to be many similarities but also obvious differences between the Nordic countries in terms of mathematical literacy performance. In all these countries the variation in performance is below the OECD average. The average level and distribution of student performances on the standardised performance scale seem to be very similar in Denmark, Iceland and Sweden. The Norwegian students' performance remains surprisingly low and, with the exception of one content area (*uncertainty*), is also below the OECD average. In contrast, Finnish students clearly outperformed their Nordic peers and were the top performers among the OECD countries. The Finnish performance profile also deviates from those of the other Nordic countries.

The high standard of Finnish students is largely due to the fact that in Finland the lower achieving students (the lowest 10 to15 percent) performed much better than the corresponding groups in other countries, also globally. This shows that a high average standard can be achieved by taking equal care of the learning of the whole age cohort. How this is achieved in practice is influenced by a whole range of factors. There are, of course, some underlying factors arising from the characteristic historical, sociological, linguistic and cultural traditions in Finland (e.g. Simola, 2005) as well as national policies regarding curricula, teacher education and mathematics instruction.

An important explanation for the high standard of mathematical literacy in Finland can be found in the development of our mathematics curriculum for the

comprehensive school (Kupari, 2005). Competency in applications and problem solving had already been approved as an essential goal of the Finnish mathematics curriculum by the beginning of 1980s. During the last 20 years this goal has gradually become well established in mathematics teaching practice. Owing to the relatively small group sizes in Finland today, it has also been possible for Finnish teachers to successfully develop teaching methods suitable for heterogeneous teaching groups and for supporting different kinds of learners. Since PISA particularly focuses on the ability of students to apply their mathematical skills and knowledge in real-life situations, the Finnish mathematics curriculum and Finnish mathematics teaching in general have given students plenty of experience in the type of problems found in PISA. Furthermore, one essential principle in the Finnish education system is a big investment in early intervention and special needs education so as to tackle learning difficulties in a timely and effective fashion.

The range of gender differences in mathematical literacy performance varies considerably across the Nordic countries. In Denmark the gap clearly favours boys, while in Iceland it favours girls in all content areas and at all performance levels. In Finland, Norway and Sweden the overall gender differences are roughly the same size (favouring boys). In Finland boys are in the majority at the both ends of the distribution on the performance scale, while girls outnumber boys at the intermediate levels. Rather a similar gender pattern can be also found in Norway and Sweden.

Other similarities can be detected among the Nordic countries regarding gender differences in mathematical literacy performance in items calling for different types of cognitive processes. In all the Nordic countries the performance differences seem to vary in the same way depending on the content area. For example, girls perform better than boys in items involving reflection in the content areas of *quantity* and *space and shape*, whereas boys do better in the areas of *change and relationships* and *uncertainty*. It is possible that the pattern observed is not only characteristic of the Nordic countries. Therefore these findings are seen as preliminary and need further analysis and verification.

In the light of our analysis the Nordic countries do not seem to constitute any one distinct group. Instead, there seem to be subgroups of countries depending on the criteria of the investigation. For instance, Denmark, Iceland and Sweden are similar in terms of their mathematics performance profiles, while Finland differs from the other Nordic countries the most in its profile, which is more like Canada or the Netherlands (cf. chapter 3 by Olsen). When we look at gender differences, Finland, Norway and Sweden form a tight subgroup. Overall, the results correspond well

with earlier PISA results for the domains of reading literacy and science (e.g. Kjärnsli & Lie, 2004; Lie & Roe, 2003). In any case, with regard to mathematics teaching in the Nordic countries, monitoring the similarities and differences, and determining the underlying reasons, is useful because this kind of reliable assessment information helps each country in its educational development and thereby ultimately promotes the learning of its students.

References

Gorard, S. & Smith, E. (2004). An international comparison of equity in education systems. *Comparative education*, 40, p. 15-28.

Husén, T. (1974). *Learning society.* London: Methuen.

Kupari, P. (2005). PISA 2003 mathematics results in Finland and in Macao: comparisons and observations. *Teacher Magazine [in Chinese]*, 12, p. 49-52.

Kjärnsli, M. & Lie, S. (2004). PISA and scientific literacy: similarities and differences between the Nordic countries. *Scandinavian Journal of Educational Research*, 48, p. 271-286.

Lie, S. & Roe, A. (2003). Unity and diversity of reading literacy profiles. In Lie, S., Linnakylä, P. & Roe, A. (eds.) *Northern Lights on PISA. Unity and diversity in the Nordic countries in PISA 2000,* p. 147-157. Oslo: University of Oslo.

OECD (2003). *The PISA 2003 framework. Mathematics, reading, science and problem solving knowledge and skills.* Paris: OECD Publications.

OECD (2004). *Learning for tomorrow's world. First results from PISA 2003.* Paris: OECD Publications.

Simola, H. (2005). The Finnish miracle of PISA: historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41, p. 455-470.

Chapter 3

A Nordic Profile of Mathematics Achievement: Myth or Reality?

Rolf V. Olsen

Abstract

This chapter presents the analysis of the so-called item-by-country interactions for the cognitive items in the domain mathematical literacy in PISA 2003. By using cluster analysis the aim is to establish whether it is reasonable to speak of a distinct Nordic profile of achievement in mathematics as this is operationalised in PISA. The analyses presented give evidence for such a Nordic profile in mathematics. However, the Nordic profile is tightly linked to the profile of the English-speaking countries. There are some implicit messages for educational policy that may be drawn from the analyses presented in this chapter. It is, for instance, important to note that the success of Finnish students in mathematical literacy in PISA is not entirely due to a distinctly different profile. To some extent Finnish students have the same relative strengths and weaknesses as their Nordic peers. This implies that overall, the Finnish students are stronger than their Nordic peers in all aspects of mathematics covered by PISA. Hence, if the data from large scale international comparative assessments are perceived as a resource for learning from others, this finding implies that detailed studies of the subject matter of the curriculum are not necessarily the way ahead.

Nordic abstract

Kan man snakke om en egen nordisk profil i prestasjoner i matematikk i PISA 2003? I dette kapitlet forsøkes dette spørsmålet besvart ved å ta utgangspunkt i at alle land har en karakteristisk relativ prestasjonsprofil på tvers av alle matematikkoppgavene i testen. Ved hjelp av klyngeanalyse vises det at de relative prestasjonsprofilene i de nordiske landene deler mange av de samme karakteristiske trekkene, og at det derfor er meningsfylt å snakke om en egen nordisk profil i matematikk i PISA 2003. Analysen viser videre at profilene i de nordiske landene også deler mange likhetstrekk med de engelskspråklige landenes profiler. Det er også flere andre tydelige klynger av land som har beslektede prestasjonsprofiler i matematikk. Det er interessant å registrere at de

finske elevene i noen grad har de samme relative styrkene og svakhetene i sine prestasjoner som elevene i de andre nordiske landene. Dette betyr at de finske elevene presterer bedre enn elever i sine naboland på de fleste matematikkoppgavene i den faglige testen, uavhengig av hva oppgaven dreier seg om. I en utdanningspolitisk sammenheng er en av årsakene til at man deltar i internasjonale undersøkelser som PISA, at man ønsker å betrakte sitt eget lands prestasjoner mot en generell internasjonal bakgrunn, og man ønsker ofte også å relatere sitt eget lands prestasjoner til utvalgte referanseland som demonstrerer hva som er mulig å oppnå. For de andre nordiske landene brukes ofte Finland som en slik referanse. Men i denne sammenhengen er det mest nærliggende å slå fast at analysen som presenteres her, viser at vi neppe kan finne en forklaring på den finske suksessen i matematikk i form av en genuin finsk tilnærming til matematikkfaget.

Introduction

This chapter studies the patterns across cognitive items in the domain mathematical literacy (sometimes referred to as 'mathematics' throughout the article) from PISA 2003. This study explores and compares what is here labelled as countries' *relative achievement profiles*. From a global viewpoint these profiles identify the relative strengths and weaknesses of countries. From a more near-sighted perspective these profiles identify, for each mathematics item, the degree to which the country performs better or worse than could be expected, given the average level of achievement for the country and the overall international difficulty of the item. Specifically, this procedure constructs the Norwegian profile by subtracting the average performance of the Norwegian students, the Swedish profile by subtracting the average performance of the Swedish students etc. Moreover, for all countries the international average difficulty level is also subtracted (for more details on the method used, see Olsen, 2005a). In this study the relative achievement profiles for all participating countries in PISA 2003 are included in order to use the full international sample as the initial reference.

From prior research on similar data it is reasonable to expect that countries with geographical, linguistic, political or economical similarities have related profiles. Of specific interest in this paper are the Nordic countries which in prior studies, to a varying degree, have been shown to have profiles across cognitive items that are relatively similar to each other. Indications of such a Nordic profile have been established in analyses of reading items from PISA 2000 (Lie & Roe, 2003) and mathematics items from the Third International Mathematics and Science Study (TIMSS 1995) (Grønmo et al., 2004; Lie et al., 1997; Zabulionis, 2001) as well as in analyses of science items from TIMSS 1995 (Angell et al., in press; Grønmo et al., 2004; Lie et al., 1997) and science items in PISA 2000 (Kjærnsli & Lie, 2004).

A Nordic profile is particularly noticeable in the analysis of items from TIMSS 1995[1], while in the analysis of PISA 2000 items the indications are weaker. It is also worth noting that a similar analysis of science items from PISA 2003 does not suggest that there is a distinct Nordic profile of relative achievement in science (Olsen, 2005b).

The study of similarities or differences between countries' profiles in mathematics in an achievement test may be done in several ways. Technically speaking, the task is to compare the achievement data for different countries, and thus identify countries with similar patterns in the data. Comparative analyses of these data may be done in a number of different ways and, consequently, different overall conclusions regarding the existence of a Nordic profile in mathematics achievement may be suggested. The reasons for such potential discrepancies between different studies all aiming at describing the Nordic profile are manifold: First of all different studies may include different kinds of data to represent a country's profile of achievement. Secondly, different studies may include different countries with which to compare the Nordic countries, and the degree of similarity is to some extent dependent on the selection of 'others' to which the Nordic data are contrasted. Thirdly, the measure of similarity and the criterion for stating that objects are similar may differ between studies.

An everyday analogy illustrates this. In comparing two people we may reach different conclusions regarding their similarity depending on the property that is measured: our conclusion if we focus on personality may differ from the conclusion when we focus on physical appearance. Furthermore, independently of the criteria used to make the comparison, it is never possible to conclude that two people are *complete* opposites or totally equal; in other words, there are degrees of similarity. Returning to the study of international assessments of students' achievement, it is evident that no pairs of countries have identical profiles, regardless of how the profile is constructed. However, when comparing pairs of countries, it is sometimes the case that some countries are more similar to each other than other pairs. The task ahead of us, therefore, in seeking to find a Nordic profile is to see whether the Nordic countries are more similar to each other than they are to other countries regarding their relative achievement profile. In another chapter Kupari & Törnroos (chapter 2) also seek to establish whether there is such a thing as a Nordic profile of mathematics achievement. Given the discussion above, it is important for the reader, in order not to be confused, to realise that the analyses presented here differ from those presented by Kupari and Törnroos regarding the type of data analysed, the selection of countries included in the analysis, and the criterion applied for

---

1. Finland did not participate in TIMSS 1995. Hence, the term Nordic refers to Denmark, Iceland, Norway and Sweden in this case.

identifying similarities and differences between countries. This also explains why the two studies reach somewhat different conclusions regarding the existence of a Nordic profile of mathematics achievement.

Method

The data input for the analyses presented in this article is a matrix of the p-values (the percentages of correct responses) on each mathematics item in the PISA 2003 cognitive test for each of the participating countries. The number of items is 84[2] and the number of countries included in the analysis is 41. The p-values across items for high performing countries will in general be relatively high compared to those for low performing countries. Similarly; the p-values for difficult items will in general be low across all countries compared to easier items. These overall patterns are of little interest when we seek to find country-specific patterns across items. The p-value matrix is therefore transformed to cancel out these general effects by subtracting the item-specific and the country-specific parts from all the p-values. These *p-value residuals* (Zabulionis, 2001) represent the achievement for a country on a specific item, beyond what can be expected from the item and country averages alone, or in other words, it is the *relative achievement profile* for each country. This and similar measures are therefore most often referred to as *item-by-country interactions*. In international studies of achievement, such as PISA, efforts are made to minimize such interactions (Adams & Wu, 2002; OECD, 2005). Higher item-by-country interactions imply increased error in the overall measurement, and even more seriously, if the interactions are systematic in any way, for instance if they are tightly related to format, the validity of cross-national inferences may be threatened.

Having calculated the p-value residuals, the issue of similarity is approached by comparing all the correlations between each of the countries' profiles. With 41 countries there are more than 800 comparisons to be made between pairs of countries. These pairwise comparisons have been analysed using hierarchical agglomerative cluster analysis (Everitt et al., 2001). Cluster analysis is a generic term for methods aiming to cluster individual cases (or variables) into larger groups which at the same time are (a) similar to the other objects within the group and/or (b) dissimilar to the objects outside the group. To put it simply, the use of cluster analysis on these relative achievement profiles should guide us in grouping countries that are similar to each other.

---

2. In the test material there were originally 85 items, but one item was subsequently excluded from further analysis.

The result of such cluster analyses are often presented as *dendrograms* (as seen in Figure 1). Dendrograms depict the overall structure in the data. They illustrate when and how countries merge to form the clusters. In the left-hand side of the dendrogram all the countries are separated, referring to the fact that all countries have unique profiles across all the mathematic items However, some countries have highly correlated profiles, and this is shown by lines merging at various distances towards the right-hand side of Figure 1. This process continues until all the countries finally are grouped together in one large group. The overall result is an illustration of the hierarchical structure with which the method organises the data. In effect, the underlying cluster structure in the data is visualised. The important decision to be made is when to stop while reading the diagram from the left to the right. If there is an interesting clustering pattern in the data this will obviously lie somewhere in between the two extremes.

However, the dendrograms may also hide some of the structure. Supplementary analyses have therefore also been executed in order to highlight the degree of internal cohesion in the proposed clusters. The mean correlation coefficient averaged over all pairwise comparisons of countries within the cluster is one indicator for this internal cohesion. Coefficient alphas[3] are calculated to study the degree to which it is meaningful to aggregate the profiles for the countries within the cluster. Furthermore, the coefficient alpha statistics makes it possible to study whether some countries should be left out of the cluster. A full account of the method used to analyse the data is given in Olsen (2005a).

Results

The dendrogram in Figure 1 presents possible cluster structures at different distances from left to right. At short distances (from the left in the dendrogram) there are some pairs and triplets of countries that are tightly related, e.g. Hong Kong/Macao or Czech Republic/Slovak Republic. This short distance illustrates the fact that the relative achievement profiles for these countries are highly correlated (correlation coefficient of around 0.70). However, in this chapter the aim is to identify larger clusters of countries, and in particular to study the potential clustering of the five Nordic countries. These more macroscopic structures are typically seen further to the right. Inspecting the diagram with this in mind, several clusters seem to be present. They are presented below with a focus on the shaded area in the upper quarter of Figure 1 where the Nordic countries are included.

---

3. Frequently such coefficients are also referred to as Cronbach's alphas. In analyses later in this chapter average profiles for clusters are calculated and used. These averages are sum scores, and thus it is appropriate to use coefficient alpha as a test of the internal consistency reliability for average profiles for a group of countries.

Figure 1 *Dendrogram of the cluster analysis of countries. The Nordic-English cluster is highlighted*
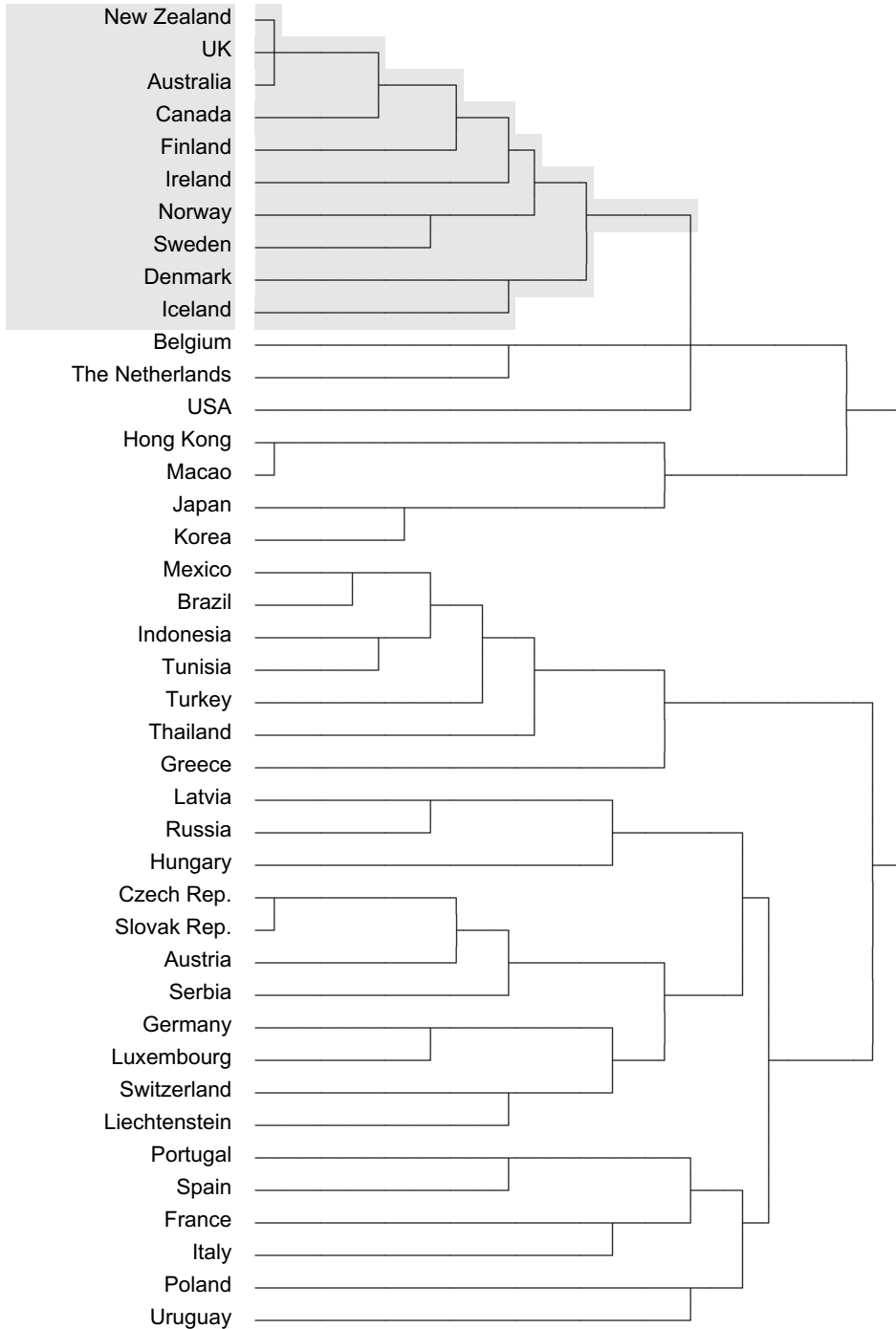
The cluster of English-speaking and Nordic countries

Based on the dendrogram and subsequent exploration of the data, the cluster in the upper quarter of Figure 1 consisting of five English-speaking countries (New Zealand, UK, Australia, Canada and Ireland) and the five Nordic countries is fairly well established, with a strong degree of internal coherence (see Table 2) and a reasonably clear degree of external isolation. Based on the dendrogram we initially considered whether the group consisting of the Nordic and the five English-speaking countries should be expanded to include Belgium, the Netherlands and the USA. The final decision not to include these three countries was mainly based on the fact that these three countries are only weakly or moderately correlated to the other countries in the cluster. Furthermore, the countries eventually included in this cluster make up the most coherent group as evaluated by the coefficient alphas. If any of the three countries entering the cluster further to the right are included the coefficient alpha decreases.

It is also worth noting that the five English-speaking countries seem to cluster tightly with each other. Finland is the Nordic country most strongly related to the English-speaking countries, while the rest of the Nordic countries join the cluster further to the right.

Table 1 *Correlations between countries' relative achievement profiles and the average cluster profiles of the Nordic and English-speaking cluster. * The correlation with the average profile is with the average for all countries excluding the country correlated with the profile. This is done in order to cancel out autocorrelation effects*

	Nordic*	English
Denmark	0.55	0.42
Finland	0.50	0.58
Iceland	0.47	0.34
Norway	0.65	0.50
Sweden	0.52	0.50

The close relationship between the Nordic and English countries is also evident from Table 1. This table shows the correlation between each Nordic country's relative achievement profile and the average profile for the five English-speaking countries, and with the average profile of the other Nordic countries. The label 'Nordic' in Table 1 thus refers to slightly different groups for each of the countries: For example, in the entry for Denmark, the correlation is between the Danish profile and the average profile for the four other Nordic countries. This ensures that these measures are not inflated, and as such the average values for the

correlations with the average Nordic profile may, for each country, be compared to the correlation with the average profile for the English-speaking cluster of countries. These correlations show that Finland is the Nordic country most closely related to the English-speaking group, while Iceland is more distantly related. All Nordic countries, except Finland, are more closely related to the average Nordic profile than to the average profile of the English-speaking countries. Of the Nordic countries the Norwegian profile appears as most closely aligned with the average Nordic profile.

Thus far we may tentatively conclude that in a comparison of all countries participating in PISA 2003 the Nordic countries are most similar to each other, except for Finland which is somewhat closer to the English-speaking countries. Nevertheless, the Finnish profile is also highly correlated with the profiles for the other Nordic countries. It is therefore reasonable to see what the structure looks like if the English-speaking countries are not included in the analysis. From the data presented in Figure 1, and Tables 1 and 2 we can reasonably expect that in the absence of the English countries, all five Nordic countries would join together in a coherent cluster.

Figure 2 *A section of the dendrogram when New Zealand, UK, Australia, Canada and Ireland are excluded from the analysis*
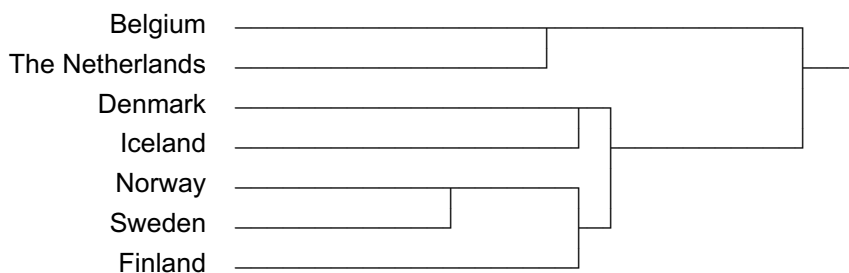

Figure 2 shows a section of the dendrogram obtained when the five English-speaking countries are excluded from the cluster analysis. In this dendrogram the Nordic countries appear as a cluster with high internal cohesion and a high degree of external isolation from the other countries. Also included in Figure 2 are the nearest neighbouring countries, the pair consisting of Belgium and the Netherlands, which can be seen to merge with the Nordic countries much further to the right. In other words, in the absence of the English-speaking countries a cluster consisting of the five Nordic countries, including Finland, is established.

As stated in the introduction, the motivation for including countries other than the Nordic group in the analysis was mainly to explore whether the Nordic countries would appear relatively similar to each other compared to all other

Table 2 *Correlations between clusters of countries. The significant correlations (p < 0.05) are shown in bold. The upper numbers in the cells in the diagonal give the coefficient alpha / average correlation within the cluster, and the lower numbers give the range of correlation coefficients within the cluster*

	N	Nordic + English	Nordic	English	East Asian	Less developed	Central European	East European	German Speaking
Nordic+ English	10	0.88 / 0.45 [0.19, 0.82]							
Nordic	5	**0.90**	0.77 / 0.40 [0.27, 0.57]						
English	5	**0.92**	**0.64**	0.88 / 0.63 [0.40, 0.82]					
East Asian	4	-0.11	0.00	-0.18	0.71 / 0.38 [0.15, 0.81]				
Less developed	6	**-0.61**	**-0.62**	**-0.49**	-0.13	0.86 / 0.51 [0.26, 0.66]			
Central European	7	**-0.47**	**-0.34**	**-0.52**	-0.16	-0.15	0.80 / 0.37 [0.11, 0.79]		
East European	6	**-0.52**	**-0.39**	**-0.55**	-0.18	-0.05	**0.73**	0.76 / 0.34 [-0.05, 0.79]	
German speaking	4	-0.22	-0.15	-0.25	-0.19	**-0.29**	**0.82**	0.34	0.72 / 0.42 [0.34-0.55]

Table 3 *The countries identified as belonging to common clusters. Observe that the three last groups are partly overlapping*

Cluster label	Countries included
Nordic + English speaking	New Zealand, UK, Australia, Canada, Finland, Ireland, Norway, Sweden, Denmark and Iceland
Nordic	Denmark, Finland, Iceland, Norway and Sweden
English speaking	Australia, Canada, Ireland, New Zealand, UK
East Asian	Hong Kong, Macao, Japan, Korea
Less developed	Mexico, Brazil, Indonesia, Tunisia, Turkey and Thailand
Central European	Czech Rep., Slovak Rep., Austria, Serbia, Germany, Luxembourg and Switzerland
East European	Latvia, Russia, Hungary, Czech Rep., Slovak Rep. and Serbia
German speaking	Germany, Austria, Switzerland and Luxembourg

countries participating in PISA. As a by-product of this analysis other clusters of countries are also revealed. The suggested clusters are identified in Table 3, but the reasons and analytical basis for the choices that have been made will not be explicitly presented. In general the procedures used to identify these other clusters are the same as those briefly illustrated above for the Nordic cluster. Observe for instance that the last three of the proposed clusters in Table 3 are not mutually exclusive. The initial group labelled as 'Central European' consisted of some East European countries and some German-speaking countries. This composite group was therefore studied further in the same way as was done for the Nordic-English cluster above: When all the German-speaking countries were excluded the East European countries clustered more clearly (except for Poland), and similarly when the East European countries were excluded from the analysis, the German-speaking countries made up one clear cluster. It is therefore reasonable to suggest that the larger, more complex cluster of several Central European countries could be replaced by two smaller and less complex structures comprising the East European and German-speaking countries

Table 2 presents the result when all countries were aggregated into the suggested clusters. In the uppermost data-containing cells running diagonally across the table the coefficient alphas and the average within-group correlation coefficients are given. These cells also give the range of the correlation coefficients between the countries included in the clusters. The minimum values represented in this range show that even if the coefficient alpha and average correlation coefficient are high, there may be individual correlation coefficients indicating that within a cluster there are two country profiles that are quite dissimilar. The other cells give the correlation coefficients between the average cluster profiles. It is clear that not only are the country profiles within the clusters relatively similar (as indicated by the coefficient alphas and average within-group correlation coefficients in the uppermost cells), but also that the average profiles across clusters are dissimilar (as indicated by the moderate to high negative correlation coefficients between the average cluster profiles).

From a Nordic perspective Table 2 confirms and establishes the degree to which the Nordic countries are on average positively related to the English-speaking countries. The minimum correlation coefficient of 0.27 is for the relationship between the Icelandic and the Swedish profiles, while the maximum coefficient of 0.57 is for the relationship between Norway and Sweden. Furthermore, Table 2 clearly shows that the Nordic profile is distinctly different from the East European countries and from the group of countries labelled as 'less developed countries'. It is also worth noting that the Nordic profile is neither positively nor negatively related to the East Asian countries,.

Discussion and possible implications

Given the results obtained from similar studies previously conducted using similar data, the proposed clusters could be expected. However, given that we might expect that the data include a large random error component, the existence of these clearly identifiable groups of countries is surprising[4]. From a measurement point of view, such fluctuations are considered to be random errors, and in the developmental phase of the test material efforts are made to reduce the influence of these item-specific patterns across countries. Even though item-by-country interactions may be perceived as a source of error in the international measurement of achievement, there is, in the case of PISA, no reason to conclude that this error source threatens the aim of international comparisons. On average, the p-value residuals in the above analyses correspond to a standard error of international measurement (Wolfe, 1999) in the range 0.5 – 1 percentage points.

In this chapter the p-value residuals have not been studied from a test perspective. Instead they have been considered as data points that give meaningful descriptions of countries' relative achievement profiles. Hence, the residual values have been assumed to contain substantial information describing the interaction between the countries and the items. At the outset, it is not reasonable to expect that all items measure the same overall trait consistently. If the fluctuations expressed by the p-value residuals are only random errors of measurement, the analysis of the profiles would probably not end up generating a systematic pattern. Clearly, though there are very systematic patterns in the data across countries. We can therefore be reasonably confident that the cluster structures in the data are reliable since this pattern fits with several other similar analyses of comparable datasets. On the other hand, what these measures indicate is still not entirely clear. Chapter 4 by Olsen & Grønmo will explore this in more depth.

The overall motivation for this work was to explore the degree to which the Nordic countries have similar relative achievement profiles across the mathematics items in PISA. The analysis suggests that the Nordic countries have highly related profiles, and these profiles are also strongly related to five of the six English-speaking countries participating in PISA. It is interesting to note that in a similar analysis of the science achievement data, the Nordic countries' relative achievement profiles were strongly related to the German-speaking countries' profiles (Olsen, 2005b). This was also the case in similar analyses of the mathematics and science items in

---

4. The way these residuals are constructed ensures that they sum to zero for each country; in other words, they are, for each country, fluctuations around the overall item-by-country expected means. The most stable components of the p-values are cancelled out when subtracting the item and country averages. It is therefore reasonable to expect that the residuals include the major part of the random variation in the p-values.

TIMSS 1995 (Grønmo et al., 2004) where a Nordic-English similarity for mathematics items and a Nordic-German similarity for science items were observed. This suggests that even if many of the same clusters of countries reappear in similar analysis based on other datasets, and in other domains, there are also domain-specific relationships between these clusters. This probably indicates that different school subjects may have different histories of policy exchange between countries.

The fact that the high-scoring Finnish students also have a relative achievement profile that is closely related to the other Nordic countries may appear surprising. One simple conclusion is that the Finnish students are better than their Nordic peers in all aspects of mathematics as defined by PISA, but relatively they have many of the same kinds of strengths and weaknesses (to be identified more closely in chapter 4). Analysis of data from international comparative studies from a policy perspective is often driven by a desire to learn from others. In this particular case the analysis indicates that we cannot easily use the item-specific information to identify formulae for success in the Finnish curriculum. Another fact not included in the results presented above, the close-to-zero correlation between the Finnish profile and the other high performing East Asian countries, tells us that there must be several different curricular recipes for success. The most likely conclusion in this quest for success factors is therefore that international differences in performance levels in achievement tests such as PISA and TIMSS are most likely related to factors other than international differences in the subject matter of the curriculum. This conclusion is also supported by the fact that across all the domains tested in PISA the same countries more or less consistently perform well. This suggests that the extent to which factors related to curricula may contribute to the understanding of high achievement they must be related to more general and overarching elements in the way the curriculum is organised or delivered, and not specific parts of it, such as the different weights put on, for instance, pure or applied mathematics.

Large-scale international studies of students' achievement are frequently criticised for being used mainly to rank countries. However, the analysis presented in this chapter has demonstrated that even if the process of test development used in PISA aimed to remove items with large item-by-country interactions, the remaining small p-value residuals may still be used to establish a clear cluster structure in countries' relative achievement profiles. This shows that achievement studies like PISA provide data that may be used to report more than the average achievement of countries. There is a fine structure within the data, and this evidently not only reflects the random errors in the measurements, but can actually be used to describe and analyse the diversity of mathematical achievement across the participating countries. It is therefore reasonable to believe that analyses such as that presented here and in chapter 4 by Olsen & Grønmo can provide valuable

information for further research in comparative education. However, in order to utilise this type of information to target specific issues we need to find ways to link these relative achievement profiles to descriptions of the policy and teaching related to the subject of mathematics across education systems.

References

Adams, R., & Wu, M. (Eds.). (2002). *PISA 2000 Technical Report*. Paris: OECD Publications.

Angell, C., Kjærnsli, M., & Lie, S. (in press). Curricular and cultural effects in patterns of students' responses to TIMSS science items. In S. J. Howie & T. Plomp (Eds.), *Contexts of learning mathematics and science: Lessons learned from TIMSS*. Lisse: Swets & Zeitlinger Publishers.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster Analysis* (4 ed.). London: Arnold.

Grønmo, L. S., Kjærnsli, M., & Lie, S. (2004). Looking for cultural and geographical factors in patterns of response to TIMSS items. In C. Papanastasiou (Ed.), *Proceedings of the IRC-2004 TIMSS* (Vol. 1, pp. 99-112). Lefkosia: Cyprus University Press.

Kjærnsli, M., & Lie, S. (2004). PISA and Scientific Literacy: similarities and differences between the Nordic countries. *Scandinavian Journal of Educational Research, 48*(3), 271-286.

Lie, S., Kjærnsli, M., & Brekke, G. (1997). *Hva i all verden skjer i realfagene? Internasjonalt lys på trettenåringers kunnskaper, holdninger og undervisning i norsk skole,*Oslo: Institutt for lærerutdanning og skoleutvikling, UiO.

Lie, S., & Roe, A. (2003). Unity and diversity of reading literacy profiles. In S. Lie, P. Linnakylä & A. Roe (Eds.), *Northern Lights on PISA* (pp. 147-157). Oslo: Department of Teacher Education and School Development, University of Oslo.

OECD. (2005). *PISA 2003: Technical Report*. Paris: OECD Publications.

Olsen, R. V. (2005a). *Achievement tests from an item perspective. An exploration of single item data from the PISA and TIMSS studies, and how such data can inform us about students' knowledge and thinking in science.* Series of dissertations submitted to the Faculty of Education, University of Oslo, No. 48. Oslo: Unipub.

Olsen, R. V. (2005b). An exploration of cluster structure in scientific literacy in PISA: Evidence for a Nordic Dimension? *Nordina, 1*(1), 81-94.

Wolfe, R. G. (1999). Measurement Obstacles to International Comparisons and the Need for Regional Design and Analysis in Mathematics Surveys. In G. Kaiser, E. Luna & I. Huntley (Eds.), *International Comparisons in Mathematics Education*. London: Falmer Press.

Zabulionis, A. (2001). Similarity of Mathematics and Science Achievement of Various Nations. *Education Policy Analysis Archives, 9*(33).

Chapter 4

What are the Characteristics of the Nordic Profile in Mathematical Literacy?

Rolf V. Olsen and Liv Sissel Grønmo

Abstract

The previous chapter dealt with the development of the relative achievement profiles for countries across all the items within the domain of mathematical literacy in PISA 2003. Analyses of the overall pattern in these profiles indicated that the Nordic countries' profiles were relatively strongly correlated. Furthermore, the analysis also revealed that other groups of countries with similar relative achievement profiles existed. In this chapter these clusters are further analysed by studying characteristics of the items. The findings reveal that the profiles of the Nordic and the English-speaking countries are mainly accounted for by variables describing what could be termed 'realistic mathematics'. This finding is discussed in relation to curricular approaches to mathematics competency and learning in the Nordic countries.

Nordic abstract

Det forrige kapitlet handlet om hvordan man kan utvikle såkalte relative prestasjons-profiler på tvers av alle oppgavene i matematikk for hvert land. Analysene av disse profilene viste at de nordiske landene i noen grad presterer relativt godt eller dårlig for de samme oppgavene. Analysene viste også at det finnes tilsvarende grad av samsvar mellom profilene for flere andre grupper av land. I dette kapitlet søker vi å forstå hva som karakteriserer de relative prestasjonsprofilene for de nordiske landene – både som én gruppe og hver for seg. Dette gjøres ved først å karakterisere hver enkelt oppgave ved hjelp av ulike klassifiseringer. Det overordnete funnet er at de nordiske landene presterer relativt sterkest på oppgaver som tester det vi kaller 'realistisk matematikk'. Resultatene diskuteres i lys av hva som har vært fremtredende mål for grunnskolematematikken i disse landene.

Introduction

Identifying possible explanations for why groups of countries have similar relative achievement profiles across mathematics items, as presented in the previous chapter, is a challenging task. Given the obvious identity of several of these clusters, or the ease with which we could label them, it is of course tempting to jump to conclusions about cultural antecedents and systemic factors that could explain the background characteristics of these groups of countries. However, it may not be so easy to derive these conclusions from the data themselves. There are few, if any, variables in the PISA data that can be used to describe the historical, social, political, economical or cultural factors that could be seen as background influences leading to links between countries in mathematics achievement profiles across several items.

The most direct approach to describing the profiles for the clusters of countries using the data available is to identify more precisely the particularities of the items defining each of the profiles. This may be done in several ways. One possibility would be to first identify a smaller group of key or defining items, in the relative achievement profiles. These items could then be investigated in detail in order to develop verbal descriptions of the profiles that go beyond the very specific context of the items. Alternatively, one could start by describing the items in more abstract or general terms, and then apply these descriptors to all the items. Of the two alternatives the latter is preferable since this approach utilises all the items in the pool. Thus, this approach has the potential to identify key aspects of the achievement profiles of the clusters that are largely independent of the actual items in the pool.

Method

The item pool in mathematics consists of 84 items. Some of the items refer to a stimulus material of some length, some don't. Some items ask the student to formulate their own answers, while some ask students to select the most appropriate from several given responses. Some of the items refer to phenomena related to everyday life, while others refer to scientific phenomena. And so on, the point being that it is possible to develop a great range of different descriptors characterising the items, and in some cases the item characteristics may be present to varying degrees. Some of the item descriptors we have chosen to develop are directly based on the PISA framework (OECD, 2003), while others were developed independently. The framework categories are accounted for in more detail in chapter 2 by Kupari and Törnroos.

The descriptors are in principle independent of the actual items present in PISA, and could have been used to describe mathematics items in general.

Table 1 *Description of some broad descriptors used to classify the mathematics items*

Variable name	Description [with variable values in brackets]
Item format	Classification of items into the two main formats, constructed responses [1] or selected responses [0].
Context four Context two	The mathematics items in PISA are classified in the framework into four different situations (personal [0], educational/occupational [1], public [2] or scientific/intra-mathematical [3]) which may be perceived as ordered according to the distance from the situation to the student ('Context four'). The context may also be classified dichotomously by separating the scientific context [1] from the other contexts set in a more real-life context [0] since this is a purer context where the mathematical aspects themselves are central ('Context two').
S & S C &R Quantity Uncertainty	The mathematics items are classified in the framework according to four phenomenological topics; Space and Shape ('S & S'), Change and Relationship ('C & R'), Quantity and Uncertainty. The four variables suggested classify the items as belonging to each of the topics or not.
Competency	The items are classified in the framework according to the main competency involved. The three competencies range from what has been termed the 'reproduction' cluster items [0] to the 'reflection' cluster items [2]. In between these two extremes is a group of items classified as the 'connections' cluster [1].
p-value	The overall international average p-value of the items.
RealMath	The degree to which an item confronts the students with a realistic problem relating to their personal lives or to citizenship.
Algebra	Classifies the items that include an explicit algebraic expression versus those not including algebraic expressions.
Calculations	Classifies the items that to a large degree require calculations.
Graphics	Classifies the items that include graphical representations of quantities.
Tables	Classifies the items that include representations of quantities in tables.
Non-continuous	Classifies the items that include information presented in a non-continuous way, e.g. graphs and tables, but also sketches or illustrations other than graphs
Complex reading	Classifies the items that mainly require the student to handle the information given, for instance by sorting relevant from irrelevant.

However, we have selected the descriptors that we initially perceived to be of particular relevance for the PISA study and for this particular item pool. A useful descriptor would be one that may be formulated so that independent evaluators of the items would agree whether an item had the characteristic or not, and furthermore, a useful item descriptor would divide the items into groups or assign a value to the items that would vary across them.

Many of the descriptors in Table 1 are very straightforward and people would easily agree whether an item has the characteristics or not. The descriptor labelled as RealMath has however been constructed in a more complex manner. The two authors of this chapter started by each independently evaluating whether the items were dealing with 'realistic mathematics', that is, mathematics perceived to be of particular relevance for everyday life or for life as a citizen. The definition of mathematical literacy in PISA highlights this aspect of mathematics, but nevertheless, the items in the pool refer to 'real life' mathematics to a varying degree. Even though according to the framework of the study all items in PISA relate to mathematics as students are supposed to need it in their daily or civic lives, the degree of relevance differs.

In the two independent evaluations it became clear that the two authors had applied slightly differing definitions or criteria for evaluating whether the items dealt with mathematics in real life settings, 'real life' referring to an item dealing with a mathematical problem or competency that is likely to be relevant to all citizens at one time or another, as opposed to items that are only likely to be relevant under very specific conditions. Even if we reconciled our definitions there were discrepancies in our evaluations of the items regarding this characteristic. One of the authors assigned a single value to each of the items in a holistic manner. The other author evaluated three slightly different aspects of the authenticity of the items: 1) whether or not the stimulus material presented had been extracted from an authentic text; 2) whether or not the item related to a realistic problem in the context supplied by the stimulus material; and 3) whether or not the underlying competency tested in the item could be considered highly relevant to 'realistic mathematics'. Except for the variable categorising the authenticity of the stimulus material, all the variables describing how realistic the items were, were highly positively correlated to each other. It was therefore decided to establish the construct RealMath as the sum of the holistic evaluation system developed by one of the authors and the system for evaluating the realism of the problem and competency components of the item developed by the other author. Constructed in this way, the RealMath variable represents 'realistic mathematics' as a broader concept than can be represented by a single dichotomous variable. The coefficient alpha for the construct was close to 0.8. This number reflects the degree of consistency with which the two authors evaluated items as being 'realistic' or not. Examples of items that were categorised as RealMath are 'Robberies' and 'Internet Relay Chat', examples of items that were

Table 2 *Correlations between average cluster residuals and the broad item descriptors. The significant coefficients (p<0 .05) are shown in bold*

	Context two	S & S	C & R	Quantity	Uncertainty	P-value	RealMath	Algebra	Calculations	Graphics	Complex reading
Nordic+English	0.14	**-0.24**	0.05	-0.07	**0.26**	**0.25**	**0.35**	**-0.28**	**-0.29**	**0.37**	0.11
Nordic	0.07	-0.15	-0.10	0.05	0.20	**0.28**	**0.32**	**-0.37**	-0.19	**0.29**	-0.05
English	0.18	**-0.27**	0.17	-0.17	**0.27**	0.17	**0.32**	-0.15	**-0.33**	**0.37**	**0.23**
East Asian	-0.19	0.11	-0.13	-0.04	0.06	-0.17	**-0.24**	0.07	0.02	-0.17	-0.15
Less developed	-0.04	0.12	-0.04	-0.19	0.11	**-0.45**	-0.15	0.16	0.05	-0.06	0.04
Central European	**-0.23**	0.19	-0.12	**0.39**	**-0.47**	0.19	-0.13	0.02	**0.24**	**-0.42**	-0.19
East European	-0.02	0.20	0.02	**0.29**	**-0.52**	0.21	**-0.31**	0.15	**0.32**	**-0.30**	**-0.23**
German speaking	-0.14	0.13	-0.07	**0.30**	**-0.36**	**0.22**	0.07	-0.11	0.10	**-0.22**	-0.02

not categorised as RealMath are 'Walking' and 'Step Pattern' (see Appendix for these items)

Results

Table 2 presents the correlations between the average cluster profiles of item residuals and the broad item descriptors presented in Table 1. Only the descriptors significantly correlated (p< 0.05) with at least one cluster profile are included. It is, for instance, noteworthy that the variable reflecting the item format is not included. In other words there seems to be no effect of format across the mathematics items and across countries. Each of the clusters may be characterised based on the information in Table 2, but here we will only consider the Nordic perspective. This means that the main goal of all comments will be to highlight typical aspects of the Nordic profile, although to some extent this includes pointing out how the profile differs from those of the other groups of countries. The relative achievement profile for the Nordic countries is clearly characterised by a

comparatively strong performance in items containing realistic mathematics. This is also true for the English-speaking countries, but is distinctly different from the profiles of all the other clusters, except the German-speaking countries where this variable does not seem to affect students' achievement. Students from the East Asian and East European clusters of countries achieved relatively lower scores in items containing realistic mathematics, and this was also the case but to a lesser extent for students from the Central European and 'less developed countries'.

Our categorisation of items in terms of their realistic mathematics content seems to provide a useful variable for describing the characteristics that distinguish the Nordic and the English speaking-countries from other clusters of countries. It seems that realistic mathematics is an important focus in school mathematics in Nordic and English-speaking countries, as opposed to the situation elsewhere. The focus on realistic mathematics in the Nordic and English-speaking countries is also typical of the PISA definition of mathematical competency, the main goal being to test students in the mathematics they may need in daily or civic life.

The performance of students from the group of Nordic countries in items including explicit algebraic expressions is relatively weak. This is consistent with findings from TIMSS 1995 (Lie *et al.* 1997) and 2003 (Grønmo *et al.* 2004). Students from the Nordic countries also perform relatively better in items including diagrams or graphs. In addition there are non-significant indications that the Nordic achievement profile is relatively strong for items relating to *Uncertainty* and weaker for items related to *Space and Shape*. These results seem consistent with the result showing a strong performance by the Nordic countries in items containing realistic mathematics: algebraic expressions are not commonly used in daily or in civic life, while in contrast knowledge about diagrams, graphs and uncertainty are likely to be useful.

Students in the Nordic countries also tend to perform relatively better in easier items, and to achieve lower scores in items requiring calculations, although this tendency is non-significant. The performance of students from the English-speaking countries in items requiring calculation is significantly low compared to other countries, while students in Central and East European countries achieve relatively good scores in these items. This may indicate that the ability to perform accurate mathematical calculations is not seen to be as important in the Nordic and the English-speaking countries as in other parts of Europe. The extent to which the increased focus in Norway over the last few decades on those aspects of mathematics needed in daily or civic life has resulted in too little attention being given to accurate calculation has been discussed in other articles based on PISA and TIMSS data (Bergem *et al.* 2005; Grønmo 2005, Grønmo & Olsen, 2006 In press).

Our analyses of the relative performances in specific items of students in different clusters of countries shows that the Nordic cluster profile has some distinct characteristics, and that this profile is closely related to that of the English-

speaking countries. These characteristics are clearly different in several ways from the relative achievement of students elsewhere, especially students in East and Central European countries and to a lesser extent also students in the East Asian and the less developed countries.

In general it is worth noting that the item descriptions taken from the framework classifications do not suggest an explanation of the Nordic profile (and only occasionally provide an explanation for the other clusters). Overall, the item descriptors developed exclusively for this analysis seem to be more successful in accounting for the cluster structure[1].

Diversity in the Nordic profile

Although the Nordic countries are very similar to each other, compared to the other participating countries, they also have different relative achievement profiles across the variables describing typical features of the items.

Table 3 *Correlations between the Nordic countries' p-value residuals and the broad item descriptors. The significant coefficients (p< 0.05) are shown in bold*

	Diversity			Unity			
	Context four	Non-continuous	Uncer-tainty	p-value	Algebra	Graphics	RealMath
DNK	**-0.22**	**0.28**	0.00	**0.37**	**-0.24**	0.05	**0.22**
FIN	0.19	0.06	0.04	0.17	**-0.24**	**0.24**	0.21
ISL	0.05	0.03	**0.27**	0.09	**-0.25**	0.17	**0.24**
NOR	0.04	0.06	**0.30**	**0.24**	**-0.29**	0.21	**0.28**
SWE	-0.01	0.10	0.07	0.19	**-0.32**	**0.37**	0.19

Table 3 presents the correlations between the item descriptors and the p-value residuals in the Nordic countries. Again, the table only includes the item descriptors that correlate significantly (p< 0.05) with at least one country profile. It has to be noted that the figures under the heading Non-continuous are the partial correlations when controlled for Graphics[2]. Thus, the numbers in this column give the

---

1. A multiple regression analysis with the average Nordic profile as the dependent variable and all the item descriptors as independent variables has been performed. The independent variables were entered in two blocks with one block representing the descriptors based on the framework categories, while the other descriptors were kept in a separate block. Independently of which block was entered first, the analysis showed that the descriptors based on the framework could account for about 10% of the variance in the average Nordic profile, while the other descriptors could account for about 30% of the variance.
2. This has been done because the variable identifying items with non-continuous stimulus material also includes all items with diagrams or graphs. This creates an autocorrelation, and by calculating the partial correlation the autocorrelation is removed in the numbers reported in the column with the heading Non-continuous.

correlation between the profiles and the occurrence of non-continuous materials other than graphical representations of quantities (e.g. tables, illustrations etc.).

What stands out in Table 3 is that it confirms the commonalities of the Nordic profile as described above: All Nordic countries show a relatively weak performance on items including explicit algebraic expressions and the Nordic profiles are similar for items assessing realistic mathematics. Even though not all of the coefficients are significant, the profiles in the Nordic countries tend to be quite similar, showing relatively good performances in items with graphical representations of quantities and, for most Nordic countries, a more pronounced preference for easier items. However, Table 3 also reveals some differences between the countries.

Danish students perform relatively better in items where the context is familiar to the student. Furthermore, Danish students do not exhibit the Nordic characteristic of performing relatively better in items including graphical representations of quantities. Instead they perform relatively better in items including non-continuous material other than graphical representations. The Danish profile is also very strongly correlated to the p-values of the items, indicating that Danish students are relatively more successful in items with low difficulty. The Finnish profile has, in general, somewhat weaker correlations with the item descriptors, but it does not deviate very clearly from the average Nordic profile. The largest deviation from the other Nordic countries is a non-significant tendency for the Finns to perform relatively better in items where the context is more remote from the student's experience. Icelandic students perform relatively well in items relating to *Uncertainty*, and the Icelandic profile is less affected by the item difficulty. Similarly the Norwegian students perform relatively strongly in items relating to *Uncertainty*. Furthermore, for Norwegian students, as for Danish students, the difficulty of the item is a significant factor. Students in both these countries have a stronger preference for easier items than, for example, students in Iceland. The Swedish profile mainly follows the average Nordic profile.

Although there are some differences between the Nordic countries, the remarkable similarities that constitute a specific Nordic profile in mathematics achievement are most obvious. It is worth reflecting on this, especially considering that Finland is a high achieving country, in contrast to the other Nordic countries. It follows from this that the specificity of the general Nordic profile is not the reason that most Nordic countries do not perform particularly well in PISA.

Discussion and possible implications

Our analyses suggests that on average the students in the Nordic cluster of countries perform relatively better on items setting realistic tasks; on items relating to graphical stimulus material; on items with overall low difficulty; and on items with no explicit algebraic expressions. Furthermore there are indications that the

Nordic profile is relatively strong for items relating to probabilities and statistics set in a realistic context (referred to in the PISA framework as *Uncertainty*), and for qualitative tasks that do not require the student to perform accurate calculations. Taken together, these findings are consistent with the assumption that the so-called 'realistic mathematics' has had a strong influence on curricula in the Nordic (and English-speaking) countries (Niss, 1996; de Lange, 1996; Gardiner, 2004).

In the Norwegian context it is reasonable to claim that an emphasis on everyday applications of mathematics has been the most important driving force underlying the changes in the curriculum over the last few decades (Alseth et al., 2003). While mathematics relating to everyday applications receives more attention today than some decades ago, the teaching of algebra and other more formalised aspects of mathematics has over the same period been reduced. This development is probably affected by the wider and longstanding discussion within mathematics education research about the role of applied mathematics in school mathematics programmes. From the mid-1980s, there has been a lot of discussion about this issue (Lange, 1996; Kilpatrick *et al.* 2005). Our experience is that, particularly in the Nordic countries and in the English-speaking countries, this has lead to applied mathematics playing a major role in school mathematics curricula. The documented strengths and weaknesses of these countries therefore probably reflect the curricular priorities in the Nordic and English-speaking countries.

Central to the discussion about the role of applied or 'realistic' mathematics has been the distinction between pure and applied mathematics. On the one hand it has been argued that all mathematics is to some extent applied mathematics (Kline, 1980), while on the other hand there are those who argue that pure mathematics exists by more or less denying any kind of usefulness (Hardy, 1967). It seems difficult to base a fruitful discussion about school mathematics on such extreme viewpoints, and this may partly be the reason that some researchers have concluded that the dichotomy between pure and applied mathematics is in itself not very fruitful (Lange, 1996), and why others use quotation marks for the terms 'pure' and 'applied' (Niss, 1999; Niss & Jensen, 2002). It has been useful to us for analysis and discussion in this article to make a distinction between pure and applied mathematics and we have used the terms without any quotation marks.

The discussion seems particularly relevant in relation to the results from PISA for the Nordic countries. The finding that the Nordic countries perform relatively well for items resonating well with the overall curriculum aims in these countries is, of course, not very surprising. However, it may be seen as unsatisfactory that all the Nordic countries, except Finland, perform at an average level in the PISA 2003 test in mathematics, even though the curriculum in these countries has emphasised teaching mathematics in real-life settings, in accordance with the definition of mathematical competency in PISA. Based on these results it seems relevant to reconsider how fruitful the strong emphasis on real-life mathematics has been, and

it seems to be relevant to ask whether this emphasis has become too dominant. More specifically, it is reasonable to suggest that we should carefully consider whether the most necessary basic skills in mathematics, e.g. skills relating to simple arithmetic, have been neglected to an extent which is also unproductive for the overarching aim of fostering the mathematics competencies needed for everyday and civic life. Mathematics in real life as *an alternative* to learning basic skills in pure mathematics may not be the best option for mathematics in school (Gardiner, 2004). It may seem to be a paradox to suggest that in order to foster a mathematical competence characterised by the ability to deal with realistic problems, we should put less emphasis on mathematics in 'an everyday mode' in the curricula. But given the findings presented here this is at least an issue that deserves a closer inspection before new curriculum plans are developed in the future.

The relation between basic skills in pure mathematics and applied mathematics has been discussed by the authors in other articles based on analyses of PISA and TIMSS data (Bergem *et al.* 2005; Grønmo, 2005; Grønmo & Olsen, 2006, in press). In these articles we have pointed out that in a Nordic country like Norway students perform relatively better in PISA than in TIMSS, while the opposite is true for Eastern European countries such as Russia. The conclusion in these articles has been that even if practice in and teaching of mathematics in real-life settings is needed in order to provide students with competencies to deal with problems in real life requiring the use of mathematics, this is not a sufficient condition to approach world class standards regarding the type of competency measured in PISA. It seems that Norwegian students' lack of elementary skills in mathematics may be one reason for their relatively low level of achievement in PISA. The results presented in this chapter lead to the same conclusion, that even though it is necessary to practice applied mathematics, a more extended mathematics programme is probably needed if the aim for all Nordic countries is to attain international standards regarding the type of competency measured in PISA.

References

Alseth, B., Breiteig, T., & Brekke, G. (2003). *Endringer og utvikling ved R97 som bakgrunn for videre planlegging og justering - matematikkfaget som kasus.* Notodden: Telemarksforskning.

Bergem, O. C., Grønmo, L. S., & Olsen, R. V. (2005). PISA 2003 og TIMSS 2003: Hva forteller disse undersøkelsene om norske elevers kunnskaper og ferdigheter i matematikk? *Norsk Pedagogisk Tidsskrift, 89*(1), pp. 31-44.

Gardiner, A. (2004): *What is mathematical literacy?* Paper presented at the ICME-10, 4.-11. July, Copenhagen.

Grønmo, L. S., Bergem, O. K., Kjærnsli, M., Lie, S., & Turmo, A. (2004). *Hva i all verden har skjedd i realfagene? Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2003.* Oslo: Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.

Grønmo, L. S. (2005). Matematikkprestasjoner i TIMSS og PISA. *Nämnaren 32*(3), pp. 5-11.

Grønmo, L. S., & Olsen, R. V. (In press 2006). Matematikkprestasjoner i TIMSS og PISA. In B. Brock-Utne & L. Bøyesen (Eds.) *Om å greie seg i utdanningssystemet i nord og sør. Innføring i internasjonal pedagogikk.* Bergen: Fagbokforlaget.

Hardy, G. H. (1967). *A Mathematician's Apology.* Cambridge: Cambridge University Press.

Kilpatrick, J., Hoyles, C., Skovsmose, O., Vallero, P. (2005). *Meaning in mathematics education.* New York: Springer.

Kline, M. (1980). *Mathematics, The Loss of Certainty.* Oxford: Oxford University Press.

Lange, J. de (1996). Using and Applying Mathematics in Education. In Bishop, A., Clements, K., Keitel, C., Kilpatrick, J., Laborde, C. (Eds) *International Handbook of Mathematics Education Part 1* (pp. 49-97). Dordrecht: Kluwer Academic Publisher.

Lie, S., Kjærnsli, M., & Brekke, G. (1997). *Hva i all verden skjer i realfagene? Internasjonalt lys på trettenåringers kunnskaper, holdninger og undervisning i norsk skole.* Oslo: Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo.

Niss, M. (1996) Goals of Mathematics Teaching. In Bishop, A., Clements, K., Keitel, C., Kilpatrick, J., Laborde, C. (Eds) *International Handbook of Mathematics Education Part 1* (pp 11-47). Dordrecht: Kluwer Academic Publisher.

Niss, M. (1999). Kompetencer og uddannelsesbeskrivelse, *Uddannelse 99*(9), pp. 21-29

Niss, M. & Jensen, T. H. (2002). Kompetencer og matematikklæring: ideer og inspiration til udvikling af matematikundervisning i Danmark. *Utdanningsstyrelsens temahefteserie* (18).

OECD (2003). *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills.* Paris: OECD Publications.

Chapter 5

'Growing up'– The Story Behind Two Items from PISA 2003

Peter Allerup, Lena Lindenskov and Peter Weng

Abstract

In this chapter we investigate two open constructed response items in 'Growing up' from PISA 2003. Our main goal is to try to make both teachers and policy-makers aware of the valuable information contained in the so-called double-digit coding used in the open constructed response items: information which in our view can be exploited immediately by teachers in their daily work in the classroom. Some results will be presented, and we will outline further possibilities for more analyses of open constructed response items.

The story of the item also serves the purpose of providing information about one of the theoretical building blocks underlying the measurement of mathematical literacy in PISA, the so-called overarching concepts.

Nordic abstract

I dette kapitel beskæftiger vi os med en opgave "Growing Up" fra PISA 2003, hvor eleverne skal give besvarelser med deres egne ord, hvilket betegnes som 'open-constructed response item'. Hensigten med artiklen er først og fremmest at gøre både undervisere og uddannelsespolitikere opmærksomme på den såkaldte 'double-digit'-kodning der foretages af disse typer spørgsmål, og den værdifulde information man kan få fra elevernes svar på denne type opgaver; information som efter vor vurdering kan udnyttes af underviserne i deres daglige undervisningspraksis. Vi analyserer nogle resultater, og vi beskriver nogle af mulighederne for videre analyser af disse 'open-constructed response items'.

Kapitlet tjener også det formål at informere eksemplarisk om hvordan man i PISA-undersøgelserne beskriver det matematiske stof vist gennem opgaven "Growing Up", som hører til under idéområdet "forandringer og sammenhænge".

Artiklen foreslår en udvidelse af PISA-designet med henblik på at forbedre mulighederne for denne form for analyser.

Mathematical literacy: a key competency

A cause of confusion in the public debate about PISA and use of the results is that the intention of PISA is not to measure achievement based on national curricula, but to measure students' mathematical literacy. This encompasses mathematical insights and processes that could form part of the competencies needed to face the complex challenges of today's and tomorrow's world. The OECD project Definition and Selection of Key Competencies (DeSeCo) aims to identify a small set of key competencies. According to DeSeCo a key competency displays three important features:

1. Contributing to valued outcomes for societies and individuals
2. Helping individuals meet important demands in a wide variety of contexts
3. Being important not just for specialists but for all individuals

Mathematical literacy as defined in PISA[1] is a key competency according to the DeSeCo definition.

All key competencies proposed by DeSeCo are grouped into three competency categories, and mathematical literacy is grouped in the first category, *the ability to use language, symbols and text interactively*. The other two competency categories are *interacting in heterogeneous groups* and *acting autonomously* (OECD, 2005. p.10).

In order to describe mathematical literacy and how it can be measured, mathematical concepts, structures and ideas are expressed in terms of the contexts in which they can be used. That is why the mathematical content in PISA 2003 is described not in classical terms as, for example, *geometry* but as *quantity; space and shape; change and relationships; and uncertainty*. Each of these categories is called an overarching idea, and together they cover the range of mathematical content typically found in other mathematics assessments and national mathematics curricula, albeit in a differently organised format.

In PISA 2003 each item is devoted to one of the four overarching ideas. In this chapter we focus on the unit in PISA 2003 labelled 'Growing up' about the heights of girls and boys of different ages. The stimulus material and mathematical issues addressed in the items are meant to assess the students' mathematical literacy in the context of *change and relationship*. Broadly speaking *change and relationship*

---

1. Mathematical literacy is defined in PISA as:
The capacity to identify, to understand and to engage in mathematics and make well-founded judgements about the role that mathematics plays, as needed for an individual's current and future life, occupational life, social life with peers and relatives, and life as a constructive, concerned and reflective citizen.

concerns the ways in which natural phenomena are manifestations of change, and how phenomena relate with each other. It might be approached through a wide range of mathematical thinking tools and methods such as mathematical equations or inequalities, equivalence, divisibility and inclusion, or by symbolic, algebraic, graphical, tabular, geometrical representations, etc.

The mathematics items in PISA are constructed as one of three basic structures: multiple choice (MC) format, 'closed constructed responses', or 'open constructed responses', where the student is expected not only to outline the idea of the solution, but is requested to describe in detail the thinking behind the way the problem has been solved. The PISA study is based on scaling procedures which require that all items can be evaluated on one scale of 'item difficulty', i.e. they share just one common property, difficulty. If the scale analyses are showing fit problems to the scale model (Allerup, 2005) or there are reasons to believe that more information can be gained by looking at the single item responses, i.e. analysing student profiles, the PISA student scores must be supplemented by single item information.

The only information used to assess the ability of the student in PISA is whether the response given is correct or incorrect – the method of achieving a right or a wrong response is not part of the assessment. During the construction of a multiple choice maths item it is necessary to think how to construct alternative response categories. In the case of 'closed constructed response' or 'open constructed response' maths items, clear and unambiguous formulations of the problem stated in the item must be provided that enable the student to come up with more than one 'reasonable' proposal for a solution to the problem. The set of possible ways of solving the problem can be systematically organised from a didactic point of view, so that a coding system for categorising the student responses can be used.

Double-digit codes: a source to information about student mathematical concepts and understanding

The two-digit codes or double-digit coding system used to score the free-response items in PISA have their origin in TIMSS: Third International Mathematics and Science Study (Lie et. al, 1996). The double-digit coding system was developed as a tool to extract more information from the student responses than could be gained if the response was simply marked as correct or incorrect. A response can be correct or incorrect in many ways, and extracting information from this variety of responses important because it makes the results of big international surveys such as PISA both meaningful and useful for teachers and thereby improves the learning experience of the students, which is the object for the assessment. A double-digit coding system that can give reliable information about both the level of correctness and the kinds

of conceptions and misconceptions connected to different approaches and strategies is a useful means of evaluating the thinking behind a student's response (Beaton, 1996).

The responses are given a double-digit code on the basis of a marking guide. The first digit indicates the correctness of the response; the second digit, the 'didactical' digit, gives information about the thinking behind the student's response. We think that the information contained in these second digits can be of value to teachers in their daily work, because knowledge of the thinking behind both correct and incorrect responses will illuminate students' approaches to different types of mathematical concepts and activities and their strategies, conceptions, and misconceptions. The didactical digit is the focus for this chapter.

The following analysis of the double-digit codes given to responses to questions in the unit 'Growing up' attempts to show how the codes can give important information. In the following section we statistically analyse, for 11 chosen OECD countries, the national distributions of the double-digit codes for responses to two questions in 'Growing up'.

Is there a Nordic profile among the group of countries?

The countries included in the analysis were Denmark and 10 other countries chosen for reasons related to the Danish National Report on PISA 2003 (Mejding, 2004). The Nordic countries were chosen for cultural reasons; Germany and The Netherlands were chosen because they are close neighbours of Denmark; Turkey was chosen because it is the country of origin of many Danish immigrants; Japan and the USA were chosen as countries from other continents with cultures of great interest generally but especially in the area of technology; Mexico was chosen because it is the country with the lowest scores.

The cognitive items in PISA represent different levels of difficulty, and contain material that is typical of different achievement levels. Some questions are relatively easy, others are more difficult. The difficulty of specific items in PISA is measured using six levels. Level 1 is the lowest level of difficulty, and level 6 the highest level of difficulty. The 'Growing up' unit contains three questions at different difficulty levels. In this chapter we will focus on the two last items in the unit (at levels 3 and 4), because they are open constructed response questions. The first question (at level 1) is a close constructed response questions, and we have not included results from that question in our analysis.

In the question M150Q2 the students are asked: According to this graph, on average, during which period in their lives are females taller than males of the same age?

The difficulty of this *change and relationships* question is level 3. At this level it is expected that, among other things, the student should be capable of tackling problems which involve multiple related representations, text and graphs, including interpretation of these formats and communication of arguments. For example, in this question the student must link and connect representations of two related graphs.

We looked at the students' responses from two different perspectives. First we focused on the most informative aspect, the different kinds of student responses indicated as double-digit codes. Then we looked at the general correctness (p-value) and the gender differences.

A first look at the responses to M150Q2 shows the following distribution:

Table 1 *Distribution of answers to M150Q2 among 11 OECD countries*

M 150Q2 Code	DNK	SWE	NOR	ISL	FIN	NLD	USA	JPN	TUR	MEX	DEU
00	7	7	9	7	6	9	14	7	20	25	7
11	29	37	25	30	26	23	43	23	29	36	23
21	55	50	55	60	62	65	37	55	26	21	57
22	3	3	5	0	3	3	3	6	6	7	3
99	6	3	6	3	3	0	3	9	19	10	10

Students' responses are coded as 00, if they are considered incorrect, as 11 if partially correct, as 21 or 22 if correct, and as 99 if question is omitted. A more detailed description of the codes are presented elsewhere

Besides looking at the code distribution within the countries, we can look at the total item scores (p-values, i.e. the percentage correct) for the countries, which is created by adding the frequency of correct responses to half the frequency of partially correct responses. The results for distribution by gender are also shown:

The immediate impression from table 1 is that the distribution of answers is similar in the Nordic countries, with a relatively small difference between Norway with the

Table 2 *The total item score to M150Q2 among 11 OECD countries*

M 150Q2	DNK	SWE	NOR	IISL	FIN	NLD	USA	JPN	TUR	MEX	DEU
Total score, p-value	74	73	73	75	80	79	61	74	48	40	72
Girls	74	75	76	78	82	78	62	78	50	42	71
Boys	73	72	69	72	78	79	60	71	47	38	74

lowest score and Finland with the highest score. However, Germany, Japan and to some extent The Netherlands have the same distribution profile, so there does not seem to be a specific Nordic profile. On the other hand the distribution in the Nordic countries is very different from that in the USA, Turkey and Mexico, so there is no general pattern of responses across all countries.

In addition to the distribution of answers we can look at the total item score across countries shown in table 2. This table shows that the scores for Denmark, Sweden, Norway and Iceland are at about the same level whereas the score for Finland is higher. This is consistent with the picture given by other general comparisons of performance in mathematics. The Netherlands is comparable to Finland while Germany has a similar profile to Denmark, Sweden, Norway and Iceland.

The question seems to be oriented towards girls because in 9 out of 11 countries the girls have higher averages than the boys. This is interesting because in general in PISA the opposite is true. Maybe the context is more appealing to girls because it specifically compares girls with boys?

In M150Q3 the following question is posed:

Explain how the graph shows that on average the growth rate for girls slows down after 12 years of age.

This question is a bit more difficult than the previous one. It is at level 4 within the overarching idea *change and relationships*. This means, among other things, that in approaching problems that involve multiple related representations and models of the real-world the student has to demonstrate a deeper understanding than is required at level 3. The student must interpret complex graphs, and read one or multiple values from graphs. A correct response to the question can be expressed in many ways.

Table 3 and 4 below correspond to table 1 and 2 above:

Table 3 *Distribution of answers to M150Q3 among 11 OECD countries*

M 150Q3 Code	DNK	SWE	NOR	ISL	FIN	NLD	USA	JPN	TUR	MEX	DEU
00	0	0	0	0	0	0	0	0	0	0	0
01	4	3	2	3	3	6	3	3	7	3	3
02	24	26	36	34	26	14	37	23	36	57	23
11	43	39	40	38	58	76	43	32	17	6	44
12	3	6	1	3	3	4	7	10	7	0	3
13	3	3	1	6	3	0	3	3	7	3	3
99	23	23	20	16	7	0	7	29	26	31	24

Table 4 *The total item score to M150Q3 among 11 countries*

M 150Q3	DNK	SWE	NOR	ISL	FIN	NLD	USA	JPN	TUR	MEX	DEU
Total score, p value	50	50	42	47	68	78	53	43	30	7	47
Girls	52	51	42	50	73	75	51	45	30	7	48
Boys	48	50	41	44	64	80	54	42	31	7	47

Tables 3 and 4 have almost the same profiles as tables 1 and 2. The distribution profiles for Denmark, Sweden, and Iceland are about the same, and the profiles for Norway and Finland show, respectively, lower and higher achievement levels, in line with the general Nordic pattern. Except for some small differences Germany has the same pattern as the Nordic countries. It is remarkable that almost all the correct responses of students in the Netherlands are coded as 11 and not one is coded as 13. Similarly in Norway all correct responses are coded as 11. More students in Iceland give responses which are coded as 13. This kind of information could be valuable in discussions about learning opportunities in different countries. The student's approach indicated by code 11 differs considerably from the approach indicated by code 13, although both are correct. The diversity we see in some countries may inspire curriculum developers and teachers in other countries. The scores show that question M150Q3 is still oriented towards girls, even though this not as clear as in question M150Q2; in 7 out of 11 countries girls have higher scores.

The above examples show the kind of information that can be extracted from double-digit codes. Further analysis of the connection between student's ability and

the single-digit codes and both types of codes combined could provide more information about students' reasoning processes.

Profiles created by double-digit codes

When responses to several items are analysed simultaneously, student profiles, i.e. response patterns based on linking response information collected from the same student across several items, can be seen. The profiles may give a consistent picture of 'correctness' or 'incorrectness' which can be used for labelling purposes. However, since the creativity of a student is unpredictable so too is the total number of possible profiles and hence labelling the profiles according to particular thinking 'strategies' could turn out to be tedious, if not impossible. Usually the labelling process is based on both empirical and theoretical techniques.

To create student profiles we need simultaneous responses to several items. Within the PISA framework this means that one needs to identify several 'open constructed response' items in the same booklet, because each student is assigned one booklet only. Only two items, M150Q02 and M150Q03, satisfy these conditions – a relatively modest result if the hope was to identify whole series of items! However, the principles behind the profile analyses given in this chapter are still valid, and both the statistical and the didactical thinking can be demonstrated using these two items

In table 5 it can be seen, for example, that 302 students fit into the specific profile '21' ∩ '01' for items M150Q02 and M150Q03. It can be seen that code 21 is the code for a correct response, while code 01 is the code for an incorrect response in which students don't mention steepness or the change in the female graph around the age of 12 years.

Table 5 *Students' responses to M150Q02 and M150Q03 (N=22043) from eleven OECD countries, showing 30 different profiles. Response types identified using PISA's double-digit (DD) coding system (see appendix to 'Growing up..').*

M150Q02	M150Q03						
codes	01	02	11	12	13	99	Total
00	115	2310	322	51	39	520	3357
11	282	3483	1642	174	197	1167	6945
21	302	2878	4018	452	337	1035	9022
22	51	454	267	31	29	161	993
99	16	190	19	2	5	1493	1725
Total	766	9315	6268	710	607	4376	22042

Table 5 raises the question of whether the frequencies presented reflect valid figures from each of the 11 countries or averages of quite different values for the individual countries? As an example, the profile '21' ∩ '13' will be analysed. In table 6 the (relative) frequencies of the profile are listed for each of the 11 countries.

Table 6 *Frequency of the profile '21' (M150Q02) ∩ '13' (M150Q03) within each country and total number of profiles. The countries are labelled by their CNT codes*

COUNTRY	Frequency specific profile	Percent specific profile	Total number of profiles
DEU	25	1.8	1381
DNK	39	3.1	1249
FIN	40	2.3	1733
ISL	44	4.3	1019
JPN	23	1.6	1413
MEX	56	0.7	8363
NLD	12	1.0	1191
NOR	7	0.6	1203
SWE	24	1.7	1390
TUR	33	2.3	1447
USA	34	2.1	1653
TOTAL	337	-	22042

It becomes clear from a simple statistical test[2] that the relative frequencies are not the same across the 11 countries. The profile '21' ∩ '13' is substantially more frequent on average in Denmark, Iceland, Finland and Turkey, but clearly less frequent in Norway.

A Nordic profile within the student response profiles

Table 6 displays one aspect of the response profiles generated by items M150Q02 and M150Q03, viz. a simple record of the frequency of one specific profile '21' (M150Q02) ∩ '13' (M150Q03). However, this does not tell us anything about the general distribution of the complete range of all 30 profiles listed in table 5. In order to analyse the consistency of the complete set of 30 profiles across countries statistical techniques other than the simple test used in table 6 must be applied[3].

---

2. Ordinary χ^2 – test analysis for contingency tables.
3. Log-linear modelling techniques are applied and results demonstrate that in the saturated model the three-way interaction parameters are insignificant.

The application of more advanced statistical techniques makes it clear that the complete 30-profile system is not consistent across all 11 countries. The weights (i.e. percentages e.g. table 7, valid for a specific country) given to the 30 profiles differ significantly from country to country. On the other hand, detailed analysis of the results indicates that the five Nordic countries form a more or less homogeneous sub-group of the 11 countries. This is confirmed by a second application of the advanced statistical technique to data from the Nordic countries only. Although Norway stands slightly apart from the other Nordic countries, the pattern displayed in table 7 for the M150Q02 vs. M150Q03 profiles remains consistent across all five Nordic countries, viz. the variation across countries is not significant. A consequence of this fact is that the probabilities that the 30 profiles will be used by Nordic students are estimated by the frequencies displayed in table 7. It should be noted that, for example, the specific profile '21' (M150Q02) ∩ '13' (M150Q03) emerges with a frequency ≈ 2.34%, in good accordance with the figures listed in table 6.

From table 7 it can be seen that the majority of students in the Nordic countries apply the response profiles '11' '21' (M150Q02) ∩ '02' '11' (M150Q03), i.e. the shaded area in table 7. In fact, nearly a third, 31%, of the students give correct responses to both items, while approximately 35% (three shaded fractions) show various types of erroneous responses to one of the items. It is, however, clear from table 7 that these students commit errors which can be classified as 'not finalising the analysis completely'.

While table 7 shows features characteristic of students in the Nordic countries, the characteristics of all other students are represented in table 8. In spite of internal variations among students in non-Nordic countries, not considered further here, table 8 shows interesting differences in comparison with the Nordic students. First of all, it is clear that only approximately 22% of the responses of non-Nordic students fit into the profile '21' (M150Q02) ∩ '11' (M150Q03), i.e. both items correct. This is significantly lower than in the Nordic countries (approximately 31%). Furthermore, it can be seen from table 8 that the average frequency of the response profile '11' (M150Q02) ∩ '11' (M150Q03) for non-Nordic students is double that of Nordic students. Interestingly, although the frequencies are small, the responses of Nordic students, on average, tend to involve 'daily language description' (22 for M150Q02) and 'not using available graphical information' (02 for M150Q03) less frequently than those of non-Nordic students.

The profile analyses conducted in this chapter can easily be extended to cases where more than two items are involved. The statistical techniques are, in fact, the same and can successfully be applied to simultaneous item responses, i.e. student

profiles. It would, however, be desirable for future analyses of response profiles in PISA, if more items of the type illustrated by M150Q02 and M150Q03 were available in *one* booklet, so that profiles can be constructed across more than just two items.

Table 7 *Frequency of the 30 profiles of responses to M150Q02 and M150Q03, using PISA's double-digit (DD) system (see Appendix: 'Growing up..'). Average estimates for five Nordic countries are presented*

M150Q02	M150Q03					
Codes	01	02	11	12	13	99
00	0.14	3.15	1.62	0.09	0.02	1.18
11	1.11	10.30	11.27	0.61	0.86	5.10
21	1.38	14.09	31.03	2.11	2.34	5.87
22	0.14	1.05	1.79	0.09	0.06	0.56
99	0.00	0.30	0.11	0.02	0.02	3.62
						100%

Table 8 *Frequency of the 30 profiles of responses to M150Q02 and M150Q03, using PISA's double-digit (DD) system (see Appendix: 'Growing up..'). Average estimates for all non-Nordic countries are presented using equal weights for the countries*

M150Q02	M150Q03					
Codes	01	02	11	12	13	99
00	0.48	4.42	1.25	0.15	0.06	2.18
11	0.96	10.45	6.51	1.13	0.42	6.24
21	1.70	13.70	22.48	4.21	1.52	8.72
22	0.21	1.61	1.22	0.21	0.09	1.13
99	0.06	0.63	0.15	0.00	0.00	8.12
						100%

Nordic democratic debate and formative assessment

Mathematical literacy is a highly contested concept, especially when it comes to formulating a detailed description that goes beyond a bare definition. In our view the concept ought to be subjected to a broader democratic discussion, and that is why in this chapter we discuss how mathematical literacy is revealed in PISA, and relate this to specific questions, because most people find concrete examples easier to understand than theoretical descriptions. It is our belief that, particularly in the Nordic countries where there is a relatively high level of participative democracy, many people, especially teachers, may be motivated and competent to engage in such discussions.

The public debate in Danish journals and newspapers on the quality and meaning of PISA that took place after the results from PISA 2003 were published in December 2004 focused on some key questions, such as What is intended to be tested?, What is actually being tested?, Are the competencies being tested really relevant today and in the future?, and Does PISA fail to test some important (mathematical) competencies? In debates in other countries, for instance in Germany, Great Britain and the USA, it has been argued that PISA does not properly assess advanced mathematical competency in intra-mathematical contexts such as solving equations, reducing algebraic expressions, making calculations with numbers and the like. That kind of criticism has not been put forward in Denmark. On the contrary, the extra-mathematical contexts chosen and reflective competencies have been critically discussed: Are PISA items sufficiently realistic to test relevant competencies?, Do PISA items differ too much from the challenges met outside school?, Could it be that students have reflective mathematical competencies that are not fully recognised by PISA? It seems as though the PISA mathematical framework with its focus on mathematical literacy rather than on pure mathematics may to some extent be more compatible with the thinking of those who engage in public discussions on educational issues in Denmark. On the other hand the Danish teachers' union has also discussed whether or not it is relevant to assess competencies according to a framework that is not derived from national curricula.

We would all probably agree that handling problems involving change and relationships is a relevant and key competency for everybody in the context of DeSeCo. But in the debate in Denmark it has been argued that the questions we analyse in this chapter do not allow students to use autonomous thinking and reflection. Using reflective knowledge and abilities results in responses that according to the marking guide should be marked as faulty (Henningsen, 2005; Lindenskov, 2005). Henningsen also remarked that the graph in M150Q2 seems to be inspired by Dutch data, which have been manipulated, and contradicts scientific facts: for instance girls from 16 to 21 grow only 2 cm, not 4. Henningsen concludes 'Personally I don't think this is in accordance with the declared purpose of working in a real-world context.' (p. 33; our translation).

We note that sometimes the construction of items involves conflict between on the one hand relevance to the real world and on the other hand the need to elicit responses from the students.

The criticism that PISA items do not acknowledge students' reflective knowledge and abilities is not empirically illuminated in this chapter but could be illuminated by further analyses of the students' written responses and by relating students'

ability to their coded and written responses. Providing teachers and policy-makers with such analyses could facilitate democratic discussions on how to unfold mathematical literacy as a key competency.

Although the epistemological interests of teachers differ from those of policy-makers, as described in table 9, both parties could benefit from analyses of the double-digit coded items.

Table 9 *Epistemological interests*

Epistemological interest	Of teachers	Of policy makers
Does PISA produce results relevant to national and international goals and visions?	Professional teacher are interested in the relevance for present national goals. Participatory citizens are also interested in international visions as a reference	Interested in relation to both national and international standards
PISA used as summative assessment – how well do the country's students perform?	A relatively low interest	Very high interest
PISA as inspiration for formative assessment – what level have the students reached, and what are their potentials	Very high interest in using and being inspired to use items as information sources among other information sources	Some interest, as they build the organisational and economic settings for formative assessment in schools

Analyses such as those we have presented and outlined in this chapter are in our view of interest to both Nordic teachers and policy-makers for practicing and managing formative assessment in schools, and analyses of this kind ought to continue to be made.

In our experience of Danish public debates about PISA, when results in the form of tables comparing the countries are published then the discussions are generally based of the tables and do not look at the information that lies behind the figures in the tables. In our opinion, the more detailed information is provided on the background of PISA and the results, the more interesting and relevant it will be for teachers to use that information both in their own thinking and teaching and for taking part in public debate. In this chapter we have tried to respond lack of information by presenting some of the reactions from the discussion in Denmark and by analysing students' responses to questions.

References

Allerup, P. (2005). PISA præstationer – målinger med skæve målestokke? *Dansk Pædagogisk Tidsskrift*, maj 2005

Beaton, A. E. et al (1996). *Mathematics achievement in the middle school years. IEA's Third International Mathematics and Science Study.* Boston: Centre for the study of testing, evaluation and educational policy, Boston College.

Henningsen, I.(2005). Et kritisk blik på opgaverne i PISA med særlig vægt på matematik. *MONA,* 1 (1), 24-43. Danmarks Pædagogiske Universitet.

Lie, S., Taylor, A., & Harmon, M. (1996). Scoring Techniques and Criteria. In M. O. Martin & D. L. Kelly (Eds.), TIMSS Technical Report (Vol. 1, pp. 7.1 - 7.16).

Lindenskov, L. (2005). Test kan ikke være virkelighed og er ikke undervisning. *MONA,* 1 (1), 82-86. Danmarks Pædagogiske Universitet.

Lindenskov, L. (2005). Hvilken indsigt kan man uddrage fra PISA - undersøgelserne i forhold til samfundets fremtidige behov. I: *PISA - undersøgelsen og det danske uddannelsessystem* (s. 165-168). København: Folketingstidende og Teknologirådet.

Lindenskov, L, Weng, P. (2005). Matematikvanskeligheder og lavt præsterende elever i Danmark. *MONA,* 1 (2), 56-75. Danmarks Pædagogiske Universitet.

Mejding, J. (red.)(2004). *PISA 2003 – Danske unge i en international sammenligning.* København: Danmarks Pædagogiske Universitets Forlag.

OECD (2005). *The definition and selection of key competencies.* Executive Summary.

Chapter 6

How Similar are We?

Similarities and Differences Between the Nordic Countries in Cognitive, Affective and Contextualised Measures in PISA 2003

Svein Lie and Marit Kjærnsli

Abstract

A series of data from PISA 2003 have been analysed from the perspective of the degree of similarity within the Nordic countries and also between them and the other participating countries. Common measures have been established for a number of factors related to learning mathematics: Achievement, learning strategies, motivation, self-efficacy, classroom and school contexts, and home background factors. Based on correlations and factor analyses, PISA constructs and variables are combined to form meaningful "meta-constructs". These are in turn used as a basis for a cluster analysis to explore some fundamental similarities and differences between the participating countries. A particular focus is to understand and describe to what extent we can speak of a Nordic type or tradition in education (particularly in mathematics). Further, similarities and differences compared to other individual countries or groups of countries are also analysed.

Nordic abstract

I dette kapitlet er en rekke data fra PISA 2003 analysert for å se i hvilken grad man kan finne likheter mellom de nordiske landene og de andre deltakerlandene. Analysene er basert på dataene som er gitt for de kognitive områdene (matematikk, naturfag og lesing), holdninger og selvregulært læring i matematikk (læringsstrategier, motivasjon og selvoppfatning), klasseromsvariable og i hvilken grad dette avhenger av hjemmebak-grunnsfaktorer. Basert på cluster- og faktoranalyser på landsnivå, er variabler og etablerte construcer satt sammen til nye meningsfulle "meta-construcer". Disse er brukt til videre analyser for å se nærmere på likheter og forskjeller mellom deltakerlandene. Det er spesielt lagt vekt på å forstå og beskrive i hvilken grad det er mulig å snakke om en nordisk tradisjon i utdanning, spesielt når det gjelder matematikk, og eventuelle likheter og forskjeller til enkeltland eller grupper av land.

Background and aim

Our aim in this chapter is to compare countries in PISA across the entire international database. The idea is to simplify the PISA 2003 international data matrix as much as possible, in order to create a general description of the information without losing too much information. Basically, the database consists of a matrix with respondents (students or school managers) from all participating countries as cases, while achievement and questionnaire items are variables. To get meaningful information out of this huge data matrix, respondents are usually grouped by countries and national subgroups according to common values of certain variables (e.g. gender or immigrant status). Likewise, variables are combined into broader constructs in order to provide reliable information on achievement (e.g. scores for mathematical literacy) or on context variables (e.g. disciplinary climate). Typically, in the international PISA reports, data for a large number of these constructs (and some stand-alone items) are reported in the form of comparison between countries. Furthermore, relations between constructs, in particular between achievement on one side and home socio-economic index or other contextual constructs on the other, play an important role in the study. Our aim in the present chapter is to go a step further and by analytical methods to simplify the data matrix along both dimensions, grouping both constructs and countries into smaller numbers of larger entities.

There have been some secondary analyses on data from PISA and also from TIMSS (Trends in Mathematics and Science Study) with the aim of grouping countries together into meaningful country groups or clusters (Zabulionis 2001, Lie & Roe 2003, Kjærnsli & Lie 2004, Grønmo et al 2004, Olsen 2005 a,b). This has been done mainly by applying cluster analysis to achievement items (item by item). Countries are thus grouped together according to similarities in cognitive relative strengths and weaknesses. The basis for the analyses is the item residuals by item and by country, i.e. how much better or worse than expected a particular country has scored on a particular item, compared to the general difficulty of the item and over-all achievement for the country. Two other chapters in this volume (chapter 3, Olsen, and chapter 4, Olsen & Grønmo) contain examples of such analyses, and their main focus is to describe similarities between the Nordic countries in the detailed profiles of mathematics achievement.

Our idea in this chapter is to apply a similar procedure to the second dimension in the data matrix as well. Achievement item data are already combined into score values for subject domains. And as pointed out above, numerous constructs have been developed from student and school questionnaire data. Detailed information on the constructs can be found in the international PISA report (OECD 2004). As mentioned above, our aim is to combine these constructs (or individual items) into meaningful larger composites, which we will call '"meta-constructs"'. The hope is

that it will be possible to provide further insight into more general differences between countries. In particular, an important aim is to try to describe some general similarities and differences between the Nordic countries in education issues.

We want to emphasise that the questionnaire data are based on students' and principals' judgments, which in many cases are highly subjective. Obviously, questions about attitudes are by purpose subjective, but it is well documented (e.g. Fischer 2004, Lie & Turmo 2005) that, for instance, Likert scales may well be open to bias depending on what responses are expected or are socially acceptable. Even more objective aspects like classroom climate and students' use of learning strategies are based on interpretations by the respondents and are therefore open to similar bias. Disorder and noise in the classroom has certainly not been measured objectively in Decibels or on the Richter scale!

Establishing 'meta-constructs'

As explained above, we start by simplifying the international database by aggregating the data matrix on the ''case dimension'' for both student and school variables to country level. Further analyses are carried out mainly at this level. There are 41 countries in total.

The next step consists of aggregating data along the ''variable dimension'', by establishing around 10 composite ''meta-constructs'' from variables and constructs. In our search for meaningful meta-constructs we have been pragmatic. The idea is to apply sequences of explorative cluster and factor analyses to aggregated country level data. This data matrix consists of both achievement data (student scores) and student and school questionnaire data (constructs and some individual items). In some cases our country level data clearly led us into meaningful meta-constructs. One typical example is *learning strategies* formed by combining the three separate learning strategies (see below). It turned out that there was a pronounced tendency in the data for countries reporting high use of one strategy to report relatively high values for the other two as well. We could then go back and check whether this tendency was also found and therefore supported at the student level, and in fact it was.

In other cases, constructs are combined on the basis of the analyses into clusters that unexpected. In these cases we have tried to give meaningful interpretations of the reality behind the clusters. But it is important for us to emphasise the rather exploratory and suggestive nature of these meta-constructs. They are meaningful only to the extent that they convey some insight into educational features of importance and provide tools for establishing meaningful country ''profiles'', as will be explained below.

Before a meta-construct was established we carried out analyses to check that the reliability (Cronbach's alpha) was acceptable. And further, we also looked carefully at correlations between constructs both between and within meta-constructs,

checking that the latter were consistently larger than the former. Each of these meta-constructs consists (with one exception) of two or more constructs, most of which are already composites of several variables. Meta-constructs were calculated for each country as averages of the national mean construct values. Through international standardization, each meta-construct was given an international mean and standard deviation of 0 and 1, respectively. It should be added that the standardization was carried out for *all* countries, not just for the OECD countries as was done for constructs in the international database. The actual numeric values therefore differ from values reported elsewhere.

An overview of how constructs (and in some cases individual items) have been combined into meta-constructs is given below. Obviously, the number of meta-constructs is somewhat arbitrary, since it is possible to go even further and look for "super-constructs" as combinations of meta-constructs. Nevertheless, the factors given in the table constitute both our basic findings and elements in our search for further results concerning differences and similarities between countries.

The following is a list of how our meta-constructs (in bold) are built up as combinations of constructs. (The construct labels used in the international database are shown in parentheses in italic for reference).

1. **Math score**
2. **Home** (socio-economical level): Parental occupational (*hisei*) and educational status (*hisced*)
3. **Teacher support**: Teacher support in math lessons (*teachs*) and student-teacher relations at school (*sturel*)
4. **Subject motivation**: Interest in (*intmat*) and instrumental motivation for (*instmot*) mathematics, in addition to motivation by competitive learning (*complrn*)
5. **Social motivation**: Attitudes towards school (*atschl*), motivation by co-operative learning (*cooplrn*),
6. **Inclusive pedagogical environment**: Sense of belonging to school (belong), students' self-efficacy in math (*mathef*), in addition to inverted values for mathematics anxiety (*anxmat*) and student ambitions (*studamb*)
7. **Learning strategies**: Control strategies (*cstrat*), elaboration strategies (*elab*) and memorisation strategies (*memor*)
8. **Accountability** (the degree to which schools are held responsible for and expected to respond to outcomes of national or local assessment): A combination of school autonomy (*schauton*) and the use of assessment for monitoring the school's own practices
9. **Time math**: Percentage of schools' total instruction time (*pcmath*) and time in minutes per week devoted to mathematics (*mmin*)

10. **Shortage of math teachers**: Reported shortage of teachers (*tcshort*) and (inverted) availability of qualified mathematics teachers (inverted *sc08Q01*)
11. **Disciplinary climate in math lessons**: This construct (*disclaim*) stands alone, since it did not easily combine with any of the others.
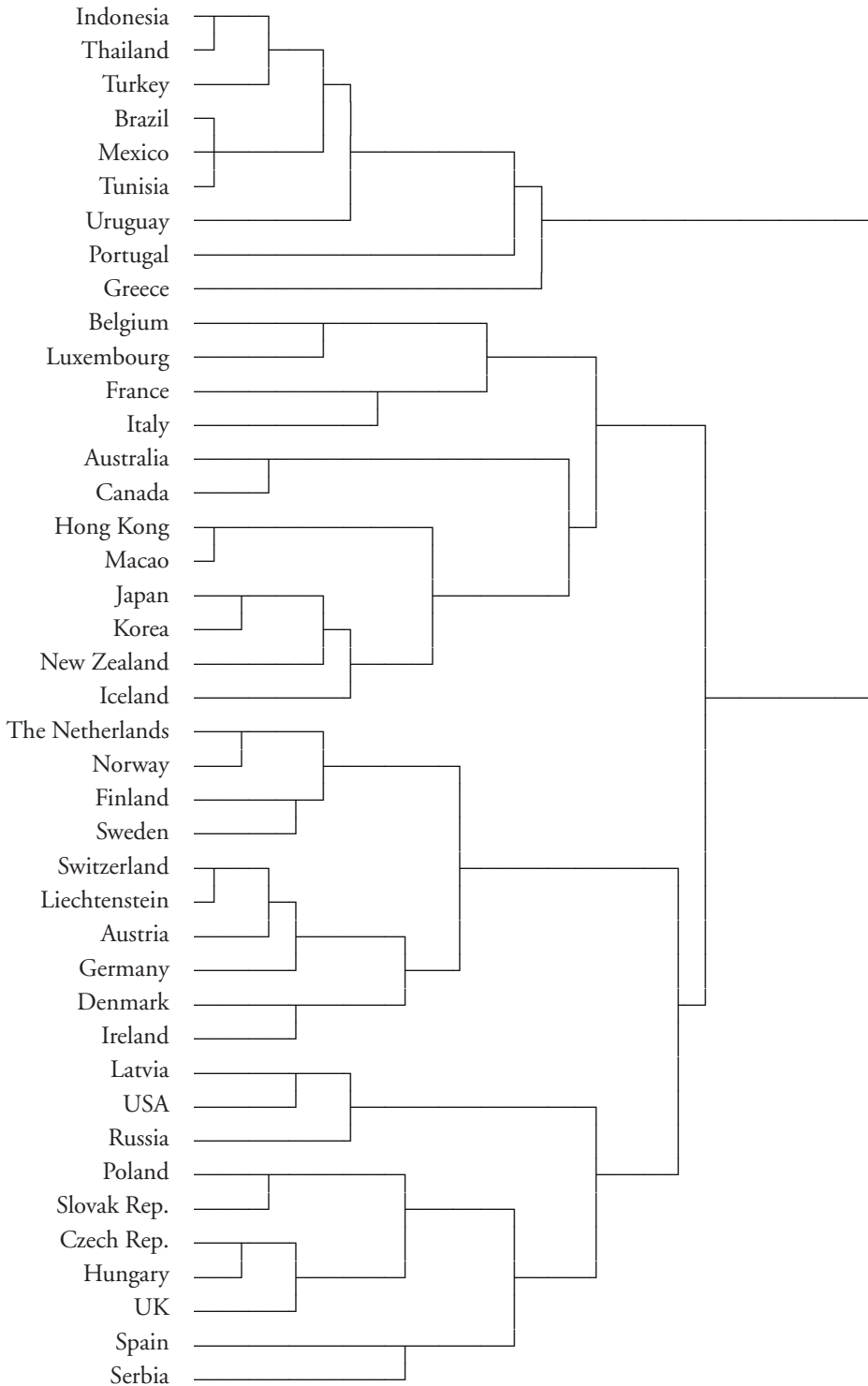
Clustering countries

The next step consisted of applying cluster analysis to see how countries group together according to similarities in their data for the meta-constructs numbered 1-11 above. Countries are clustered according to similarities in what is going on in math classrooms as well as students' attitudinal and cognitive responses thereto.

Cluster analysis is a method which allows us to systematically investigate similarities and differences between countries in the data for the actual meta-constructs. The idea is to quantify the similarities between each pair of countries by, for example, applying ordinary (Pearson) correlation coefficients. There are many alternative measures of similarities, e.g. squares of differences between the two countries summed over all meta-constructs at hand. Strong similarities between countries mean that there is a tendency for them to report high and low values for the same meta-constructs, i.e. the countries have similar patterns from variable to variable. On the other hand, consistently low or high values for a country do not influence correlations.

The analysis starts by combining the two countries with the highest similarity into one group. At the next step two other countries are clustered together, or another country is combined with a group that has already formed, depending on which similarity is the largest. Once a group is formed, its group mean for each meta-construct will then be used to calculate its similarities with individual countries or other groups. The process goes on until all countries have finally been combined into one large group.

Figure 1 is a dendrogram, which depicts how countries are clustered into increasingly larger groups going from left (high similarity) to right (low similarity). Countries with similar data tend to appear as neighbours in the country list and combine vertically relatively ''early'' (to the left) in the dendrogram. A long line of arrows before another country merges with a group reflects what is often referred to as high external isolation for the group. As mentioned above, there are many possible measures of similarity other than correlations that could be applied to the data (see for instance Olsen 2005a). By applying a series of relevant but slightly different criteria, some features appear across the versions, and we will focus on some of these in further analyses. On the other hand, there are also differences between the various versions, so one should be careful not to pay too much attention to the detail. We want to emphasise that our discussion in the following section does not relate solely to figure 1.

Figure 1 *Dendrogram of countries based on correlations of meta-constructs*

Based on figure 1 (and also on a number of alternative versions) we have suggested meaningful country clusters for further analysis. In order to do this we applied the following guiding criteria: At least four and less than eight countries should be in each group. Further, the calculated reliability (as measured by Cronbach's alpha) should be above 0.75, and each member of the group should contribute positively to alpha. And most importantly, each grouping should be conceptually meaningful in the sense that the group can be labeled in a way that provides some explanation of which countries are included and which ones are not.

Table 1 displays the proposed country groups together with measures of reliability and average (as well as minimum and maximum) correlation coefficients between countries within the group. As can be seen from the table, the groups do generally meet the criteria above. In our view there are remarkable patterns in the meta-construct data that lead to these meaningful country groups. However, there are some important comments to make at the start. Firstly, we are primarily concerned with outcomes for the actual Nordic countries. The group that for convenience is labeled ''Nordic'' differs in important respects from the actual Nordic group. More will be said about this group later. Secondly, the ''English-speaking'' group does not readily fit our criteria, since the four selected countries do not show a strong tendency to cluster early in figure 1, as they do in reading, mathematical and science achievement. It is also worthwhile noting that the European English-speaking countries Great Britain and Ireland do not fit into this group. Thirdly, for the ''East Asian'' group there is a strong tendency towards two strong pairs of countries: Japan–Korea and Hong Kong–Macao, respectively. This is no surprise, given the fact that the latter two are not countries, but semi-independent provinces within the same country, China. Finally, in spite of their close linkages in figure 1, group 7 countries (France, Italy, Belgium and Luxembourg) are not clustered very strongly, as is seen both from the low alpha and the low minimum correlation (between Belgium and France). Even labeling this group is difficult; ''French'' is used for convenience.

It should also be emphasised that 11 of the 41 countries, among them Iceland and Denmark, tend to remain isolated or to form pairs between two clusters, or they tend to combine with countries which conceptually do not seem to belong to the same group. Therefore, according to the criteria, these 11 countries have not been included in table 1.

Table 2 displays the characteristics for each of the seven groups of countries in table 1. The standardised values are shown for all meta-constructs. The extreme absolute values shown in bold indicate the most pronounced characteristic features.

Table 1 *Groups of countries with labels, reliability and correlation measures. 30 of the 41 countries appear in the groups*

Country group	Countries	Cronbach's alpha	Average and range of correlations
1. Less developed	Brazil, Mexico, Indonesia, Thailand, Tunisia, Turkey	0.94	0.74 (0.52 – 0.97)
2. "Nordic"	Finland, Netherlands, Norway, Sweden	0.93	0.76 (0.65 – 0.84)
3. English speaking	Australia, Canada, New Zealand, USA	0.76	0.50 (027 – 0.80)
4. East Europe	Poland, Slovak rep., Czech rep., Hungary	0.86	0.65 (0.46 – 0.82)
5. East Asia	Hong Kong, Macao, Japan, Korea	0.84	0.61 (0.32 – 0.91)
6. German speaking	Switzerland, Liechtenstein, Austria, Germany	0.95	0.82 (0.73 – 0.93)
7. "French"	France, Belgium, Luxembourg, Italy	0.72	0.46 (0.14 – 0.79)

From table 2 it can clearly be seen that each group has a distinct pattern that is different from that of the others. Group 1 (Less developed countries) is a particularly interesting case, since most of the values are extreme. *Achievement* score and *home background* are both very low. More surprising probably are the very high values for *Teacher support*, *Subject motivation*, *Learning strategies* and *Social motivation*. Group 5 (Eastern Asia) also stands out as having many extreme values, in particular for the same factors as group 1, but in most cases with the opposite sign,. The four European Groups (2, 4, 6 and 7) have relatively more in common, and the same is true of groups 3 and 5. Accordingly, the data provide evidence for the fact that the seven country groups can conceptually be ordered into the following three ''meta-groups'' of countries:

A. Group 1, Less developed countries
B. Groups 3 and 5 combined, English speaking+ East Asia
C. Groups 2, 4 and 6 combined, Europe

Table 2 *Average scores for all meta-constructs for the seven groups of countries. Absolute values above 1.0 are shown in bold*

	1 Less developed	2 "Nordic"	3 English speaking	4 East Europe	5 East Asia	6 German speaking	7 "French"
Math score	**-1.9**	.7	.6	.3	**1.0**	.6	.3
Home	**-1.7**	**1.0**	**1.0**	.3	-.6	.2	.3
Teacher support	**1.7**	-.3	.7	-.9	-.8	-.4	**-1.0**
Subject motivation	**1.9**	-.7	.4	-.5	-.8	-.7	-.7
Social motivation	**1.4**	-.6	.4	-.3	**-1.9**	-.1	-.4
Inclusive ped env	**-1.1**	.7	.1	.4	**-1.4**	**1.9**	.0
Learning strategies	**1.8**	-.8	.3	.2	**-1.2**	-.8	-.5
Accountability	.3	.1	.7	**1.1**	.1	**-1.2**	-.8
Time math	.7	**-1.3**	**1.0**	-.7	**1.3**	-.6	.0
Shortage teachers	**1.3**	-.4	.2	-.6	-.5	-.6	.5
Disciplinary climate	-.6	-.9	-.2	.1	**1.0**	**1.1**	-.7

The order of the factors indicates that A and C lie at each end of the spectrum, with B somewhere in between. This over-arching structure can also be seen directly from figure 1.

It is very interesting to compare our findings so far with the patterns that emerge when countries are clustered based on scores for individual achievement items. In the chapter by Olsen (in this volume), groups of countries are established which are essentially the same as those presented here. It appears remarkable to us that that the patterns are so similar, given the fact that they are based on totally different types of data. Together the two sets of analyses mutually reinforce each other in the sense that there seem to be common educational features among groups of countries which may be based on deep underlying traits linked to historical and philosophical traditions, geographic patterns and linguistic influences.

How similar are we?

As discussed above, the analysis did not reveal a distinct cluster of all the Nordic countries. Iceland and Denmark did not seem to sit comfortably in the same cluster as the other Nordic countries. Let us therefore take a look at how each of the Nordic countries fits with the other country groups. Table 3 gives the correlations for all meta-constructs between the Nordic countries and each of the seven country groups. For each country the highest correlation is shown in bold in

the table. The strong linkage between the three members of the ''Nordic'' group is certainly no surprise, since this feature has been built into the applied method. As can be seen from the table, Denmark correlates relatively weakly with the group labelled as ''Nordic'', whereas the linkage to its more southern neighbours (the 'German' group) is stronger. From figure 1 one can also see the strong similarity with Ireland. Iceland, on the other hand, tends to be most similar to the (non-European) English-speaking group, in particular to New Zealand (see figure 1). It seems as though on broader educational issues each of these two Nordic countries have ''drifted'' somewhat away from their Nordic neighbours, Denmark taking on a more ''continental'' profile and Iceland showing some similarity with ''overseas'' countries. The concept of a particular ''Nordic'' type of educational setting seems so far to be somewhat challenged by our results. To analyse this point further we need to investigate the influence on these correlation patterns of the various individual factors (meta-constructs). In particular we want to ascertain which factors challenge the homogeneity within the Nordic group of countries.

Table 3 *Correlations between the Nordic countries and each of the seven country groups. The highest correlation for each country is bolded*

	1 Less developed	2 ''Nordic''	3 English speaking	4 East Europe	5 East Asia	6 German speaking	7 ''French''
Denmark	-0.23	0.24	0.19	-0.42	-0.18	0.49	0.13
Finland	-0.70	0.91	0.23	0.44	-0.12	0.41	0.22
Iceland	-0.33	0.41	0.70	0.09	0.28	-0.13	0.38
Norway	-0.58	0.88	0.28	0.27	-0.29	0.31	0.64
Sweden	-0.55	0.88	0.08	0.49	-0.23	0.41	-0.03

A detailed comparison between all the Nordic countries of the data for the meta-constructs can be seen in table 4. In the table, the meta-constructs have been sorted by similarities, i.e. the factors with the smallest spread of values among the Nordic countries appear towards the top of the table. Based on the table we will summarise similarities and differences between the five Nordic countries in the following section.

First of all, there are distinct common features for all Nordic countries. These include in particular positive values for all countries for *Home* background, *Inclusive pedagogic environment* and *Math score*. Similarly, all the Nordic countries have negative values. for *Learning strategies* and *Disciplinary climate*. The same almost applies for *Social motivation,* except for Denmark's positive, but low value.

Table 4 *Data for meta-contructs for each of the Nordic countries. The meta-constructs have been sorted by the magnitude of the standard deviation for the distribution between the Nordic countries*

	Denmark	Finland	Iceland	Norway	Sweden
Learning strategies	-.6	-.8	-.4	-.8	-.4
Home	.6	.9	.9	1.3	.8
Math score	.5	1.1	.6	.2	.4
Inclusive ped env	.9	.2	.7	.7	1.2
Social motivation	.2	-.3	-.8	-.6	-.6
Disciplinary climate	-.5	-.9	-.9	-1.4	-.4
Teacher support	.6	-.1	.2	-.6	.6
Subject motivation	.7	-.8	.3	-.6	-.2
Shortage teachers	-1.0	-1.2	.3	.2	.3
Accountability	-1.8	-.2	.8	-.5	.7
Time math	.3	-1.5	1.1	-1.0	-1.1

A remarkable lack of similarity among the Nordic countries appears in the three last rows in table 4. Iceland and/or Denmark have mostly very different values from the other Nordic countries for all three factors. These are the main factors that have led to the low Nordic cohesion in the cluster analyses presented above. Two of these factors are related toavailability of relevant math teachers and time spent on mathematics lessons, and the third is related to the degree of *accountability* of schools. It is notable that these three factors are school factors from the school questionnaires, in contrast to the other factors on the list of meta-constructs above. It is also relevant to note that all three are somewhat variable factors, in the sense that they relate to educational policy factors which is established so to say on a day to day basis. For the less variable factors such as *educational climate*, *learning strategies*, and *students' attitudinal and cognitive responses*, the situation is similar across the Nordic countries as a whole. This can be demonstrated by calculating correlations between all five Nordic countries excluding these three constructs. With a range from 0.57 to 0.86 and mean of 0.75, the adjusted correlations would then have allowed us to conclude that the actual Nordic group of countries does fulfil the requirements for being included as a group in table 1.

Considering the other Nordic countries, the situation in Norway has raised national concerns in many respects. In spite of having a strong home background, Norwegian students tend to score lower in mathematics than their Nordic peers. The Norwegian student data show particularly low values for motivational factors as well as *learning strategies* and *disciplinary climate*. Below-average instruction time combined with an above-average shortage of math teachers add to the list of concerns. Finland's situation is an interesting case. The very high achievement score

does not provide any easy explanation based on the present data alone. Noticeable are the below-average values for nearly all school-related meta-constructs. The low value for *Shortage of teachers* is, however, clearly an advantage. According to Välijärvi et al. 2002 the availability of teachers, and in particular very gifted and well educated teachers, is one of the most important factors behind the success of Finnish students in PISA. For the other factors the data seem to be so similar to that for Sweden and Norway in particular that, contrary to reality, one would also have expected these countries to have similar achievement results to Finland.

Conclusion

The present analyses have provided a method for creating a very simplified and condensed version of the full PISA database. We have used this opportunity to investigate very basic similarities and differences between the Nordic countries. In addition, we have also constructed other groups of countries according to their particular characteristic features. Similar methods have been applied in the chapters by Olsen and Olsen & Grønmo in this volume, but their analyses are based on data from individual achievement items. The general similarity of the outcomes of the two approaches provides mutual support for the approaches and adds credibility to the findings. Our focus has been on mathematics since this domain was emphasised in PISA 2003, but similar country groupings tend to emerge across a number of cluster analyses based on different types of data and subject domain (e.g. Lie & Roe 2003, Kjærnsli & Lie 2004, Grønmo et al. 2004, Olsen 2005b). Thus, there are reasons to believe that the basic findings in the present chapter are substantially independent of subject domain.

When the PISA database is analysed from an international perspective, the Nordic countries stand out as a group with some common characteristics. We will end this chapter by highlighting some notable common concerns about Nordic education, as demonstrated by our analyses (table 4). Firstly, Nordic students seem to have a low repertoire of effective *learning strategies*, and this will obviously be a problem to an extent, depending on the degree to which instruction is centred on students' self-dependent learning activities. The somewhat problematic *disciplinary climate* also appears as a cause for concern and possible improvement. For the low values for *social motivation* combined with high values for *inclusive pedagogical environment,* we offer a common interpretation: Students tend to experience an inclusive, supporting, and not very demanding atmosphere. They would, however, have benefited from somewhat more focus on learning outcomes and less on co-operative activities. It should be noted, however, that the data in table 4 and the high Finnish mathematics score suggest that this interpretation of data is less relevant for Finland than for the other Nordic countries.

References

Fischer, R. (2004). Standardization to account for cross-cultural response bias. A classification of Score Adjustment Procedures and Review of Research in JCCP. Journal of Cross-Cultural Psychology, 35 (3), pp. 263-382.

Grønmo, L. S., Kjærnsli, M., & Lie, S. (2004). Looking for cultural and geographical factors in patterns of response to TIMSS items. In C. Papanastasiou (Ed.), *Proceedings of the IRC-2004 TIMSS* (Vol. 1, pp. 99-112). Lefkosia: Cyprus University Press.

Kjærnsli, M., & Lie, S. (2004). PISA and Scientific Literacy: similarities and differences between the Nordic countries. *Scandinavian Journal of Educational Research, 48*(3), 271-286.

Lie, S., & Roe, A. (2003). Unity and diversity of reading literacy profiles. In S. Lie, P. Linnakylä & A. Roe (Eds.), *Northern Lights on PISA* (pp. 147-157): Department of Teacher Education and School Development, University of Oslo.

Lie, S. & Turmo, A. (2005). Cross-Country Comparability of Students' Self-Reports. Paper to PISA Technical Advisory Group Meeting; 09.05.2005 - 11.05.2005

OECD (2004). *Learning for tomorrow's world. First results from PISA 2003.* Organisation for Economic Co-Operation and Development.

Olsen, R. V. (2005a). *Achievement tests from an item perspective.* Unpublished PhD, University of Oslo, Oslo.

Olsen, R. V. (2005b). An exploration of cluster structure in scientific literacy in PISA: Evidence for a Nordic Dimension? *Nordina, 1*(1), 81-94.

Olsen, R.V. ( ). *A Nordic profile of Mathematics Achievement: Myths or reality?* This volume

Olsen, R.V. & Grønmo, L.S. ( ). *What is characteristic of the Nordic profiles in mathematical literacy?* This volume

Välijärvi, J., Linnakylä, P., Kupari, P., Reinimainen, P., & Arffman, I. (2002). *The Finnish success in PISA - And some reasons behind it.* Institute for Educational Research, University of Jyväskylä

Zabulionis, A. (2001). Similarity of Mathematics and Science Achievement of Various Nations. *Education Policy Analysis Archives, 9*(33).

Chapter 7

Affective Factors and their Relation to the Mathematical Literacy Performance of Students in the Nordic Countries

Jukka Törnroos, Ingmar Ingemansson, Astrid Pettersson and Pekka Kupari

Abstract

The article discusses students' self-concept, interest, motivation, and anxiety in mathematics and the connections between these affective factors and students' mathematical literacy performance in the Nordic countries in the light of the PISA assessment. The results show that students' attitudes towards mathematics vary between the Nordic countries, but the connections to performance are quite similar. The connections between the affective factors and performance were stronger in the Nordic countries than in the OECD countries on average. The results imply that special attention should be paid to the attitudes of the weaker students in the Nordic countries, for example, by strengthening their self-concept in mathematics. This could lead to a substantial increase in their mathematical literacy performance.

Nordic abstracts

Artikkelissa tarkastellaan Pohjoismaiden oppilaiden itsetuntoa, kiinnostusta, motivaatiota sekä ahdistuneisuutta matematiikkaa kohtaan sekä näiden tekijöiden yhteyksiä oppilaiden matemaattiseen osaamiseen PISA-arvioinnissa. Tulosten mukaan oppilaiden tuntemukset matematiikkaa kohtaan vaihtelevat eri Pohjoismaiden välillä, mutta yhteydet suorituksiin ovat melko samankaltaiset. Matematiikkatuntemusten ja suoritusten välinen yhteys oli Pohjoismaissa voimakkaampi kuin OECD-maissa keskimäärin. Tulosten valossa erityistä huomiota tulisi kiinnittää heikoimpien oppilaiden matematiikkatuntemuksiin Pohjoismaissa. Tämä voisi johtaa heidän matemaattisen osaamisensa huomattavaan parantumiseen.

Artikeln handlar om Nordiska elevers själuppfattning, intresse, motivation och ängslan i matematik och dessa affektiva faktorers samband med elevers matematiska kunnande i PISA provet. Resultaten visar att elevers attityder mot matematik varierar mellan de Nordiska länderna men attitydernas samband med prestation är ganska lika.

Sambanden mellan de affektiva faktorerna och prestation var starkare i de Nordiska länderna än genomsnittligt i OECD länderna. Resultaten tyder på att man borde fästa avseende till, till exempel, de svaga elevers självuppfattning i matematik i de Nordiska länderna. Det kunde leda till märkbar förbättring i deras matematiska kunnande.

Introduction

In discussions about the effects of education, the cognitive outcomes of schooling often receive most attention. The cognitive outcomes, however, represent just one facet of the results of education. Another important facet is the affective outcomes of schooling. Therefore in PISA 2003, which focused on the mathematical literacy of 15 year olds, the answers of the following questions were also considered important: How interested are students in studying mathematics? How do students see themselves as learners in mathematics? Do the students feel stressed when studying mathematics? What kind of learning strategies do the students use when studying mathematics? All of the aspects highlighted by the previous questions were addressed in the student questionnaire answered by the students who participated in PISA 2003. How the students answered these questions and how their answers related to their performance are the main questions studied in this article.

Background and research questions

The present study takes a closer look at four affective factors in mathematics explored in PISA 2003. These factors are students' *interest in and enjoyment of mathematics*, *instrumental motivation in mathematics*, *self-concept in mathematics*, and *anxiety in mathematics*. In the following section we will give a brief overview of these four factors based on the PISA 2003 framework presented in the international PISA report (OECD, 2004, pp. 109-158). This international report also includes further references for those interested in reading more about the background theories of PISA.

Within the framework of PISA 2003, the factors analysed in this chapter are seen as characteristics of students' self-regulated learning skills. Self-regulated learning is considered in PISA to involve motivation to learn and ability to select appropriate goals and strategies for learning. One main assumption made in PISA is that these factors have a positive relationship with students' performance, and this assumption is based on empirical evidence (see OECD, 2004, p. 113).

The factors characterising self-regulated learning are interrelated but each of them represents a specific construct with a specific definition. Both *interest in and enjoyment of mathematics* and *instrumental motivation in mathematics* describe students' motivation to learn mathematics (OECD, 2004, p. 115). The difference between these two factors is that the former is related to internal characteristics, while the latter is related to external rewards such as good job prospects with the help of mathematics. Both of these constructs have previously been found to be connected to student performance. This connection was also clear in the results presented in the PISA 2003 main report. Students' *self-concept in mathematics* describes students' beliefs about their own mathematical competencies (OECD, 2004, p. 115). These beliefs have an influence on the goals students set themselves, on the choice of learning strategies used, and on the final learning outcomes. The strong connection between *self-concept* and performance was also shown in the international PISA 2003 report. *Anxiety in mathematics* could be treated as an additional factor related to attitudes, but in PISA 2003 it was, as before, considered to be a distinct factor related to students' emotions (OECD, 2004, p. 116). *Anxiety in mathematics* is negatively associated with performance, which was also clearly illustrated by the first PISA 2003 results.

There were three main reasons why the factors described above were chosen for deeper analysis. First, all four factors describe the fulfilment of important educational goals as such. Second, when analysed individually all four factors were strongly related to student performance in the international report. Finally, the first results presented in the international report also showed that for all four factors there were clear gender differences favouring boys. This is also important because gender differences in attitudes and self-concept are reflected in students' future career choices. In the Nordic countries, as well as in the OECD countries on average, a greater proportion of boys than girls choose to study mathematics on a higher level (OECD, 2004, p. 155).

With these reasons in mind, we focus on the following questions in this article:

1. Regarding the four affective factors studied, are there differences between the Nordic countries or can we recognise a 'Nordic profile'?
2. What are the relationships between the affective factors and mathematical literacy performance in the Nordic countries?

These questions will be studied from several different perspectives. We will begin by presenting the questions related to the different factors and how students answered those questions. We will then show how the affective factors are connected

to student performance. This will be done first by presenting results for each factor separately and then we will show results for some hierarchical linear models, which show how all these factors together are connected to student performance. The final analyses will also include discussion of the gender differences related to the affective factors. In all of the analyses we will show results for all Nordic countries and the average for all OECD countries.

The analyses make use of two different measures of performance in mathematical literacy: The results of PISA 2003 have been presented on a points scale with an OECD average of 500 points and a standard deviation of 100 points (OECD, 2004). This means that about two-thirds of the students in the OECD countries have a score between 400 and 600 points. Furthermore, six proficiency levels (Levels 1 through 6, in ascending order) have been determined to allow for a more precise description of students' performance. Each proficiency level corresponds to about 60 points on the mathematical literacy scale.

Students' attitudes towards mathematics

Tables 1-4 show how students in the Nordic countries responded to the statements related to each of the affective factors explored in the study. Although the present article concentrates on the general patterns seen in the results, many of the individual results seemed quite interesting: For example, the Danish students had a high self-concept in mathematics compared with the other Nordic students and the OECD average (table 3). However, when we look at the last item measuring self-concept (In my mathematics class, I understand even the most difficult work), we see that only Norway had a lower proportion of students agreeing with this statement than Denmark. These kinds of unexpected response patterns were one reason for us to include all the statements and their results in this article.

From the results in Tables 1-4 it is difficult to see any clear Nordic pattern in the students' answers. In particular, the results for students' interest in and enjoyment of mathematics (table 1) and self-concept in mathematics (table 3) varied widely across the Nordic countries. Danish students were very interested in mathematics and their self-concept in mathematics was also very high. At the other end of the scale, students in Finland had a relatively low interest in mathematics and the self-concept of Norwegian students in mathematics was also relatively low.

Two features, however, seemed to be characteristic of all Nordic countries: Most obviously, students in the Nordic countries were less anxious about mathematics than students in the OECD countries on average (table 4). Indeed Sweden and

Table 1 *Students' interest in and enjoyment of mathematics in the Nordic countries*

	Proportion of students agreeing with statements (%)					
	Denmark	Finland	Iceland	Norway	Sweden	OECD
I enjoy reading about mathematics.	48	18	33	26	49	31
I look forward to my mathematics lessons.	47	20	24	29	30	31
I do mathematics because I enjoy it.	59	25	38	34	35	38
I am interested in the things I learn in mathematics.	65	45	49	50	53	53

Table 2 *Students' instrumental motivation in mathematics in the Nordic countries*

	Proportion of students agreeing with statements (%)					
	Denmark	Finland	Iceland	Norway	Sweden	OECD
Making an effort in mathematics is worth it because it will help me in the work that I want to do later.	91	73	83	82	71	75
Learning mathematics is important because it will help me with the subject that I want to study further on in school.	88	87	85	82	86	78
Mathematics is an important subject for me because I need it for what I want to study later on.	75	74	79	75	67	66
I will learn many things in mathematics that will help me get a job.	83	76	78	73	73	70

Table 3 *Students' self-concept in mathematics in the Nordic countries*

	Proportion of students agreeing with statements (%)					
	Denmark	Finland	Iceland	Norway	Sweden	OECD
I am just not good at mathematics.	30	40	46	45	34	42
I get good marks in mathematics.	70	56	55	48	59	57
I learn mathematics quickly.	60	54	55	47	60	51
I have always believed that mathematics is one of my best subjects.	48	33	41	31	31	35
In my mathematics class, I understand even the most difficult work.	34	38	39	30	44	33

Table 4 *Students' anxiety in mathematics in the Nordic countries*

	Proportion of students agreeing with statements (%)					
	Denmark	Finland	Iceland	Norway	Sweden	OECD
I often worry that it will be difficult for me in mathematics classes.	34	50	50	47	32	57
I get very tense when I have to do mathematics homework.	26	7	19	37	14	29
I get very nervous doing mathematics problems.	15	15	17	20	11	29
I feel helpless when doing a mathematics problem.	17	26	28	31	17	29
I worry that I will get poor marks in mathematics.	41	51	59	58	46	59

Denmark were the countries where students felt the least anxiety about mathematics of all PISA 2003 countries. In addition a fairly high instrumental motivation in mathematics seemed to be a common feature of students in the Nordic countries (table 2). This applied particularly to students in Denmark and Iceland, but in the other Nordic countries students' instrumental motivation in mathematics was also slightly higher than the OECD average.

Relationships between individual affective factors and performance in mathematical literacy

In the next phase of the analysis, we looked at the connections between students' performance and the affective factors. This was done by calculating averages for students at each proficiency level defined in PISA 2003 and also for students who did not attain even the lowest proficiency level (Level 1). In the following figures results for this group of students are recorded as Level 0.

Figure 1 *Interest in and enjoyment of mathematics at different proficiency levels in the Nordic countries*
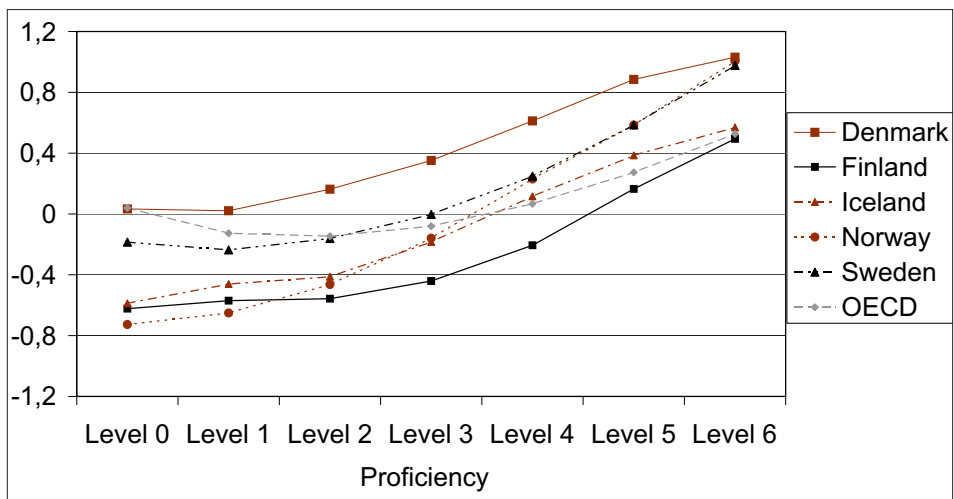
In this analysis of the relationships between the individual affective factors and students' performance, the Nordic countries seemed to have more in common than in the first analysis which concerned students' actual attitudes and feelings towards mathematics. Figures 1-4 show clearly that the curves of the Nordic countries all had quite similar shapes and these shapes differed clearly from the shapes of the OECD average curves: The Nordic curves had a gentle slope at lower proficiency levels but were steeper at higher proficiency levels. In contrast, the OECD curve had a very gentle slope at all levels and in the cases of *interest in and enjoyment of*

Figure 2 *Instrumental motivation at different proficiency levels in the Nordic countries*
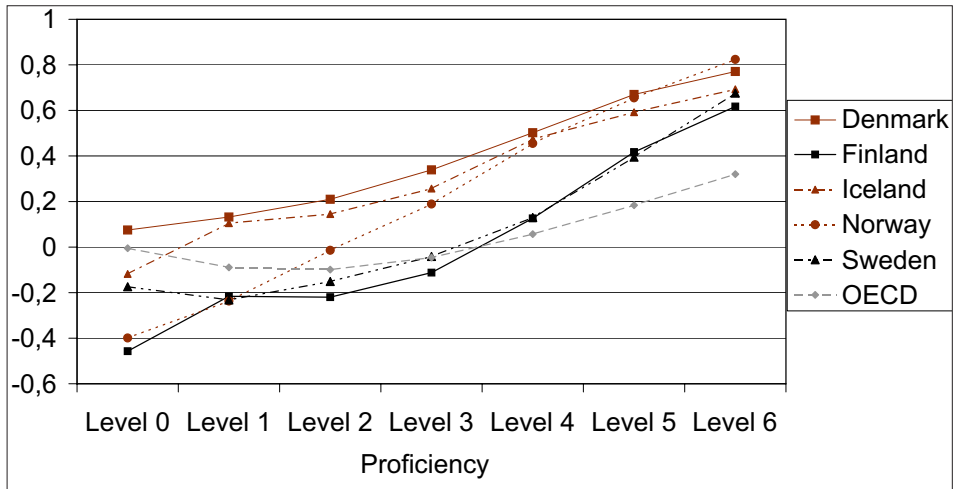
mathematics and *instrumental motivation in mathematics* the students at Level 0 had values even higher than students at Level 2.

It was particularly noteworthy that the shapes of the Nordic curves were very similar even though the actual values of the factors differed. For example, in the case of *instrumental motivation* (figure 2) the curves of Denmark and Iceland run about 0.2-0.3 scale points above the curves of Finland and Sweden, but their shapes were still very similar. Norway's curve differed a little from the other Nordic curves but its shape still seemed more like the other Nordic curves than the very flat OECD curve.

Figure 3 *Self-concept in mathematics at different proficiency levels in the Nordic countries*
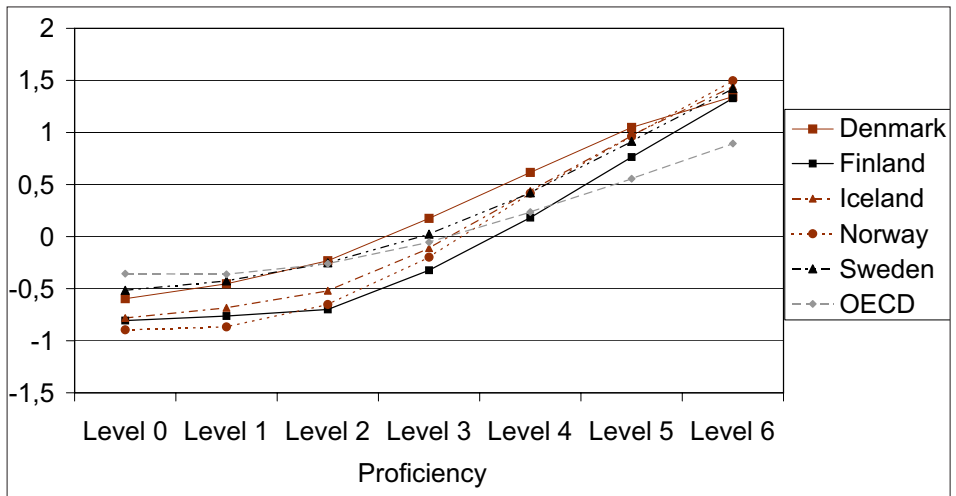
The characteristics of the Nordic curves were perhaps most explicitly seen in the curves for *students' self-concept in mathematics* (figure 3): All of the Nordic curves were steeper than the OECD average curve and they even seem to converge at higher proficiency levels.

Although the general shapes of the curves were very similar in the Nordic countries, some country-specific characteristics could be seen. In Norway the differences between students at Level 0 and Level 6 tended to be somewhat greater than in the other Nordic countries. For example, the Norwegian students at Level 0 had the lowest self-concept in mathematics of all Nordic students, but at the same time Norwegian students at Level 6 had the highest self-concept of all Nordic students. Consequently, the Norwegian curves were steeper than the curves for other Nordic countries.

Figure 4 *Anxiety in mathematics at different proficiency levels in the Nordic countries*
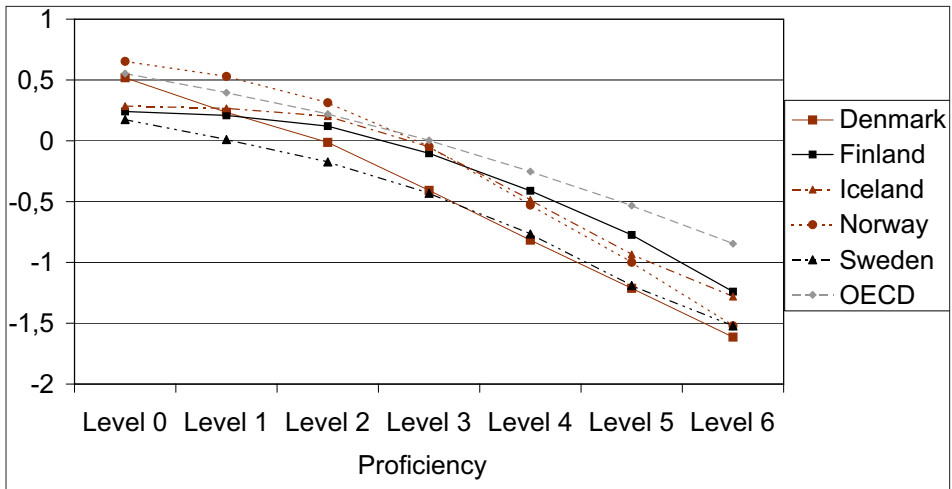
For the Finnish curves typically the differences between the three or four lowest levels were quite small and therefore the first part of the curves was very flat. As a result, in three out of the four factors studied here, the Finnish students had the lowest Level 2 values of the Nordic countries. The curve for *anxiety in mathematics* was an exception in this respect, but the first part of the curve was also quite flat.

When we looked at girls' and boys' attitudes towards mathematics at different proficiency levels, the results were very clear: boys had more positive attitudes towards mathematics than girls across the scale (although the difference was not always statistically significant). Within the Nordic countries there were only two instances when this was not the case: In Finland there was no difference between girls' and boys' *instrumental motivation* at Level 0 and in Iceland at Level 6 girls were more interested in mathematics than boys.

In addition to the general observation that boys have more positive views of mathematics than girls, one clear country-specific pattern was found. In Iceland the difference favouring boys was greatest at the lower proficiency levels, whereas at the higher levels and especially at Level 6 the gender differences were small and in the case of *interest in mathematics* girls had a more positive view of mathematics than boys, as already mentioned. No such clear patterns could be recognised for the results for the other Nordic countries.

An application of modelling: The combined connections of the affective factors

In the last phase of the present study we used hierarchical linear models (HLM) to explore further the connections between the affective factors and students' performance in mathematical literacy. These models show how strongly each of the four affective factors explored in this study were connected to students' performance, when they were studied in combination and not individually. In addition to the affective factors, gender was also used as a variable in the models. A benefit of using HLM was that it allowed us to take into account the structure of the school data: Students are clustered in schools and in the case of the OECD model schools are further clustered in countries. If this structure is not taken into account, the results could be incorrect because, for example, the results of students from the same school tend to be more similar than the results of students from different schools (the so-called intra-school correlation) (Malin, 2005).

Because of the structure of the data all factors were studied at two levels in the Nordic countries: School level contextual factors were computed by averaging the student level values of the original factors. In the OECD model the country-level was used as an additional third level and factors for this level were computed by averaging the school level factor values.

The model descriptions in Table 5 show all factors that remained statistically significant in the country or OECD average analyses. The model descriptions include only student and school level factors. This means that for the OECD model none of the country-level factors were statistically significant predictors of student achievement. When the factor coefficients are studied, it should be borne in mind that if other variables are added to the model, the coefficients for the present factors would most likely change.

In any case, according to the descriptions the affective factors were quite similarly connected to performance in the Nordic countries. At the student level, *self-concept in mathematics* had the strongest connection with performance. An increase of one

point on the self-concept scale was connected to an increase of about 40 points in mathematics performance in the Nordic countries. The effects of gender, interest and enjoyment, and anxiety were clearly smaller, and in Denmark and Iceland instrumental motivation was not statistically significantly related to students' performance. Surprisingly, interest in and enjoyment of mathematics were negatively connected with performance and the coefficient for boys was also negative.

Table 5 *Hierarchical linear models for the Nordic countries and OECD countries on average*

	Denmark	Finland	Iceland	Norway	Sweden	OECD
Intercept (factor values 0)	517.7	550.9	528.4	504.6	514.7	516.5
School level factors						
Gender (proportion of boys)						-42.4
Interest and enjoyment						-27.6
Instrumental motivation	-30.2					
Self-concept	26.6					20.3
Anxiety	-33.2	27.5		-14.7	-30.3	-49.4
Student level factors						
Gender (student is a boy)	-6.9	-15.7	-25.3	-16.8	-11.5	4.9
Interest and enjoyment	-13.8	-10.3	-8.0	-4.1	-11.0	-7.7
Instrumental motivation		5.0		8.3	6.0	3.1
Self-concept	40.3	44.7	40.9	36.2	40.7	27.0
Anxiety	-19.7	-10.3	-7.4	-14.7	-16.7	-12.4
Variation explained (%):						
Country level (OECD model)						27.8
School level	28.7	45.6	27.8	52.8	31.4	26.7
Student level	33.7	37.9	31.5	36.8	29.7	34.3
Total	33.0	38.3	31.4	37.9	29.9	31.5
	Reported effects are statistically significant with $p<0.05$.					

At the school level *anxiety in mathematics* in particular was statistically significantly connected with performance in the Nordic countries. Instrumental motivation and self-concept were also statistically significant predictors of performance at the school level in Denmark and Finland, respectively.

These simple models explained about one-third of the total variance (variation in the students' results) in the Nordic countries, the proportion explained varying from 29.9% in Sweden to 38.3% in Finland. Proportions of this size are quite typical for this kind of model (Kupari & Törnroos, 2004). However, the

proportions of school-level variance explained were surprisingly high in Finland and Norway.

How should the coefficients of the model be interpreted? When these kinds of model descriptions are presented the interpretation of the results is at least as important as the description of the model. For these models in particular two results need to be discussed: Gender differences and the role of interest in and enjoyment of mathematics. The gender coefficients (boys compared to girls) of the Nordic models are all negative. This does not mean that boys performed worse than girls in PISA mathematics tasks. In fact, boys had better results than girls in all Nordic countries except in Iceland. What these coefficients do mean is that if a girl and a boy have similar values for the affective factors, the boy has usually performed worse than the girl. Because girls reported, for example, lower self-concepts and interest in mathematics than boys in the Nordic countries, one way to improve girls' performance could be to strengthen these aspects of girls' attitudes to mathematics.

The coefficients for *interest in and enjoyment of mathematics* were indeed unexpected. *Interest* itself was positively associated with students' performance in all Nordic countries (figure 1), but when it was combined with *self-concept* in the models, the coefficients became negative. This may imply that the self-concept index in PISA also accounts, to a great extent, for the positive association of interest and enjoyment with performance. The negative coefficients for *interest and enjoyment* in the model may be caused by, for example, the fact that many students with a high self-concept and good results are not interested in mathematics. And, at the other end of the scale some students are very interested in mathematics but have great problems with learning it.

At the school level, anxiety was the factor that was most often associated with students' performance. This means that the anxiety felt towards mathematics varies between different schools and this variation is reflected in students' mathematics results. In Denmark, Norway and Sweden the effect was negative, as expected. That is, the more students felt anxiety at school the lower were their results. In Finland, however, the effect was reversed. Students' results were higher in schools where they felt more anxiety. This unforeseen result may be due to a combined effect of anxiety and self-concept at the school level, similar to the combined effect of interest and self-concept at the student level previously discussed. A similar unforeseen effect is seen in the Danish results, where instrumental motivation was negatively associated with performance at the school level.

Because the OECD average model had no statistically significant country-level factors, it looked quite similar to the Nordic models. The fact that no country-level

factors are sustained in the final models is in itself an interesting result. It means that in this model the differences between the OECD countries were less important than the differences between schools and students within the individual countries.

Some differences between the OECD model and the Nordic models are worth discussing. In the OECD average model, students' self-concept in mathematics was not as strong a predictor of performance at the student level as in the Nordic countries. It is also interesting that in the OECD model, boys had better results than girls even though their affective values were the same. In the OECD model gender also had an interesting effect at the school level: According to the model, students in classes with more girls got better results than students in classes with more boys.

In the OECD model, interest and enjoyment had a negative effect on performance both at student and school levels. This seems to validate the findings from the Nordic models but the reasons behind these findings need further research. In the OECD model, anxiety had a noticeably stronger negative effect on performance at the school level than in the Nordic countries. When all OECD countries were considered, the variation in anxiety at the school level was much greater than in the Nordic countries, which were all among the countries where students reported the least anxiety towards mathematics (OECD, 2004). In the OECD model, this increase in variation was also reflected in performance at the school level.

Discussion of the results and recommendations

Do the Nordic countries have a common profile with regard to affective factors? Based on this analysis, one could answer both yes and no. If we focus on the actual factors there is definitely variation between the Nordic countries and we can, for example, state that Danish students have the most positive views of mathematics of all Nordic students. However, when we focus on connections between the affective factors and students' performance, we can certainly talk about a 'Nordic' pattern. Compared with the OECD average the affective factors were more strongly connected with students' performance in the Nordic countries. This was seen both when connections were studied separately for each factor and when their combined effects were studied through the statistical model. Of course, there may also be other countries with characteristics similar to the Nordic countries. Nevertheless, the present study shows that the Nordic countries are moderately similar, as far as the relationships between the affective factors and student performance are concerned.

The strong connection between the affective factors and performance also helps us to identify some weaknesses in the Nordic education system. The results show

clearly that students with low performance scores have also, on average, low interest, low self-concept, etc. How could we help these students achieve positive experiences with mathematics? Although PISA highlights the connection between positive attitudes and high performance, the causal relationships between attitudes and performance are usually seen as reciprocal (e.g., Skaalvik & Valas, 1999). This leads us to suggest that students with negative attitudes should be guided in their encounters with mathematics so that they see the fascination and importance of mathematics in our world. The encounters should give these students feelings of success in performing mathematical tasks and thus give them confidence in their abilities and potential to learn mathematics. If we accept the underlying assumption of the PISA results, an increase in the confidence and interest in mathematics of these students could eventually also lead to an increase in their performance levels.

These ideas can also be applied to the gender-based results. Girls clearly showed lower levels of interest and self-concept in mathematics than boys, and they also reported more anxiety towards mathematics. The results of the statistical model, on the other hand, clearly showed that if girls had same attitudes towards mathematics as boys, they could actually achieve better results. So it seems that to close the performance gap between boys and girls in Denmark, Finland, Norway and Sweden, attention also needs to be paid to girls' attitudes towards mathematics. In Iceland the gender-based results seem more confusing in this respect (boys have more positive attitudes, but girls have higher scores) and this situation will be studied more thoroughly in chapter 14 by Ólafsson, Halldórsson and Björnsson in this book.

References

Kupari, P. & Törnroos, J. 2004. Matematiikan osaaminen peruskoulussa kansainvälisten arviointitutkimusten valossa. [Mathematics achievement in the comprehensive school according to international assessment studies]. In P. Räsänen, P. Kupari, T. Ahonen & P. Malinen, *Matematiikka – Näkökulmia opettamiseen ja oppimiseen* [Mathematics – Perspectives on teaching and learning]. Jyväskylä, Finland: Niilo Mäki Instituutti, pp. 138-169.

Malin, A. 2005. *School differences and inequities in educational outcomes. PISA 2000 results of reading literacy in Finland.* Jyväskylä, Finland: Institute for Educational Research, University of Jyväskylä.

OECD. 2004. *Learning for tomorrow's world. First results from PISA 2003.* Paris: Author.

Skaalvik, E. M. & Valas, H. 1999. Relations among achievement, self-concept, and motivation in mathematics and language arts: a longitudinal study. *Journal of experimental education, 67(2), 135-150.*

Chapter 8

Learning Strategies and Mathematical Achievement in the Nordic Countries

Are Turmo and Therese Nerheim Hopfenbeck

Abstract

What can we learn from the students' self-reports on learning strategies in PISA? In this article, we investigate the Nordic patterns found regarding students' self-reported use of learning strategies. We also explore whether any gender differences in the student self-reports can be found. Finally, we study differences and similarities between the Nordic countries regarding the relationship between the level of mathematical literacy and the students' self-reported use of learning strategies.

Nordic abstract

Hva kan vi lære av elevenes selvrapporterte bruk av læringsstrategier i PISA? I denne artikkelen studerer vi mønstre i de nordiske dataene om læringsstrategier. Vi undersøker spesielt om det er kjønnsforskjeller i egenrapporteringene i de nordiske landene. Videre studerer vi forskjeller og likheter mellom de nordiske land når det gjelder sammenhengene mellom matematikkskåre og vektlegging av læringsstrategiene. Våre analyser viser at det gjenstår en rekke ubesvarte spørsmål når det gjelder i hvor stor grad elevenes selvrapporteringer avspeiler faktisk bruk av læringsstrategier. De empiriske resultatene må derfor tolkes med varsomhet. Validering av instrumentene ved hjelp av andre metoder og tilnærminger er nødvendig.

Learning strategies in PISA

According to Chamot (1999) more than a hundred learning strategies have been identified in the research literature. The PISA study focuses on memorisation/rehearsal strategies, elaboration strategies and control strategies. *Memorisation strategies* such as different rehearsal or repetition techniques are used by the students to remember certain parts of the subject. *Elaboration strategies* are used by students to connect new information to knowledge they already possess, and are

often associated with students' efforts to gain a deeper understanding of the content. *Control strategies* reflect the learner's ability to monitor, plan and regulate the learning process, and they are therefore also called meta-cognitive strategies. The student questionnaire in PISA contains learning strategy items which are adapted from other well-known strategy inventories (Baumert et al. 1998; OECD, 2005; Turmo, 2005). Even though there are several different definitions of learning strategies depending on the researcher's theoretical orientations, most researchers predict a positive empirical relationship between the use of learning strategies and the learner's achievement level in school subjects. It is assumed that successful learners plan, monitor and regulate their own learning and use appropriate strategies in the learning process (Chamot, 1999; Weinstein, 1988; Zimmerman, 1990).

In PISA 2000, learning strategies were measured globally. This means that the students were asked about their learning strategies in general, i.e. across different subjects, subject sub-domains and contexts (Knain & Turmo 2003). According to Ramsden (1992), approaches to learning can be studied globally because of students' tendencies to use the same approaches to school work. However, the validity and relevance of such measures have been questioned (Baumert et al. 2000; Samuelstuen, 2005). In PISA 2003, learning strategies were measured in relation to the subject of mathematics, which was the major domain in this cycle.

Results at construct level

In PISA, all constructs are scaled by applying 0.00 as the OECD mean value and 1.00 as the standard deviation at student level (OECD, 2005). The results in figure 1 show that all the Nordic countries have mean values below the OECD mean value (0.00) for the strategy *Memorisation/rehearsal*. Boys report that they use this strategy more than girls in all the Nordic countries. However, the differences are greatest in Norway and Denmark. Further analysis shows that the strongest correlation between this strategy and score in mathematical literacy among the Nordic countries is found in Norway (0.26). In Norway, the correlation is somewhat stronger for girls (0.31) compared to boys (0.21). In Iceland, no empirical relationship between this strategy and the mathematical literacy score is found, while weak positive relationships are found in Sweden (0.14), Denmark (0.10), and Finland (0.08). As in Norway, the relationship is somewhat stronger for girls (0.19) than for boys (0.11) in Sweden.

The results in figure 2 show that Denmark has a mean construct value marginally above the OECD mean for Elaboration strategies, while the other Nordic countries

Figure 1 *Memorisation/rehearsal strategies in mathematics: Construct values*
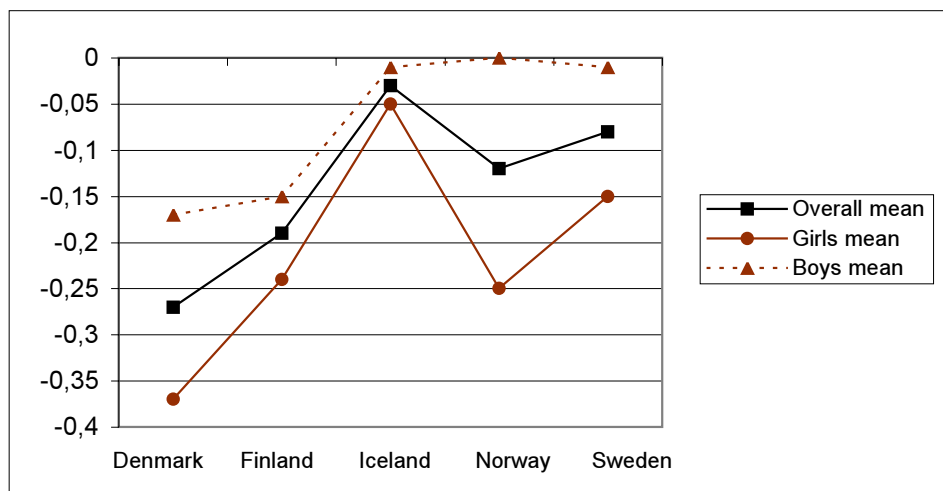
Figure 2 *Elaboration strategies in mathematics: Construct values*
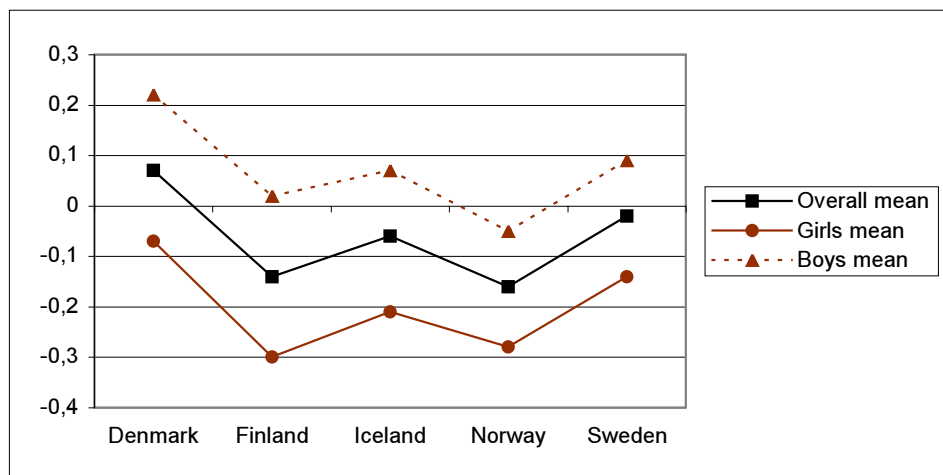
have values somewhat below this mean value. In all the Nordic countries boys report that they emphasise this strategy more than girls. Further analysis shows that the strongest relationship between elaboration strategies and score in mathematical literacy among the Nordic countries is found in Finland (correlation: 0.18). Again, no relationship is found in Iceland, while weak positive correlations can be found in Norway (0.10), Sweden (0.09), and Denmark (0.09).

Figure 3 *Control strategies in mathematics: Construct values*
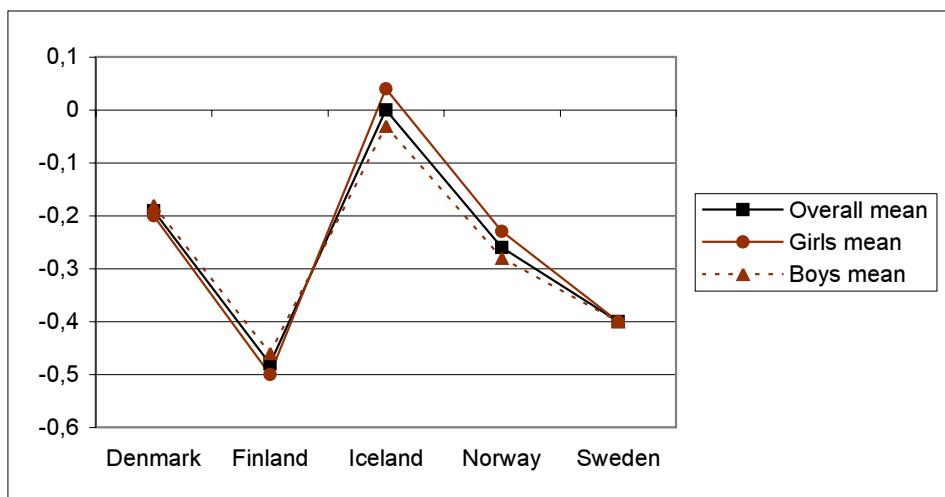
The results in figure 3 show that Iceland has the same mean value for control strategies as the OECD mean, while the other countries have values below this mean. Finland has the lowest value among the Nordic countries. There are only insignificant differences between the mean values of boys and girls in all the Nordic countries. Further analysis show that the strongest empirical relationship between control strategies and mathematical literacy is found in Norway (correlation: 0.17). No empirical relationship is found in Sweden, while weak positive correlations are found in Finland (0.11), Iceland (0.05), and Denmark (0.04).

Results at the item level

So far we have studied the results at construct level. But what are the results underlying these construct values in the Nordic countries? In the following section we will study the results at the single item level, starting with the four items included in the construct 'memorisation/rehearsal strategies'.

The results in figure 4 show that relatively few students in the Nordic countries agree with the statement about going over problems so often that they feel as if they could solve them in their sleep. The percentages vary from 19% in Denmark to 33% in Sweden. However, many students agree that they try to remember every step in a procedure for learning mathematics. The largest percentage is found for Norway (79%) while the percentage in Sweden is clearly lower (61%).

Figure 4 *Memorisation/rehearsal strategies in mathematics: Percentages of the students who 'Strongly agree' or 'Agree'*
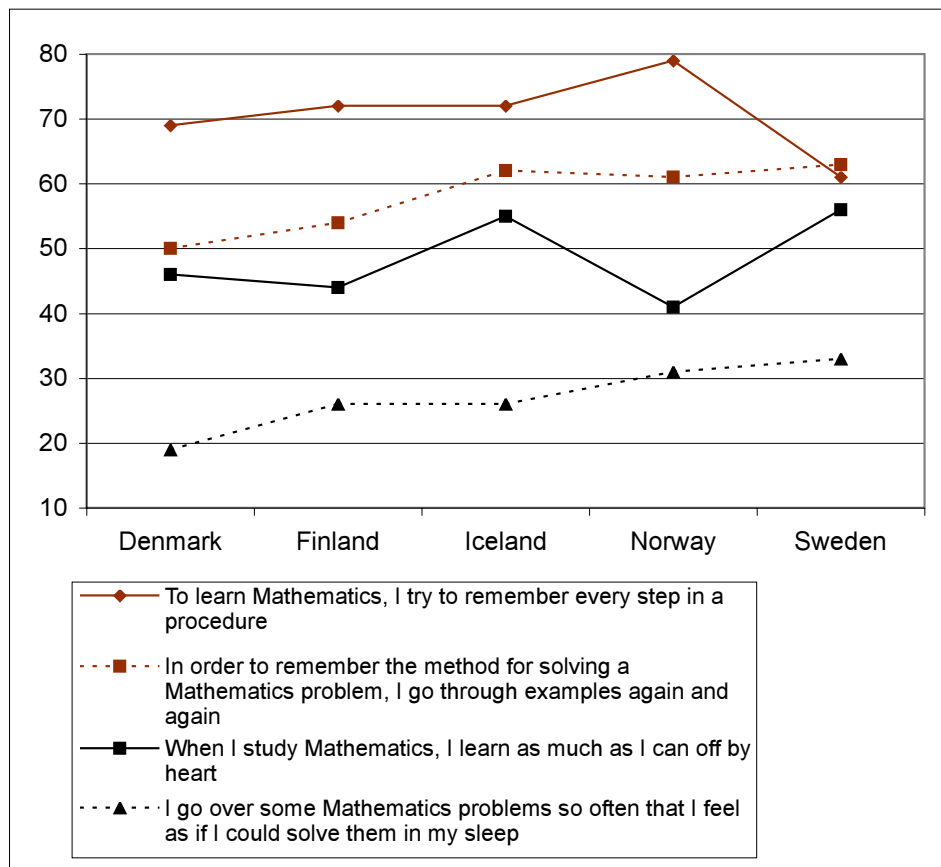
The results in figure 5 show some striking features concerning the items included in the construct 'elaboration strategies'. It is interesting to note that only 27% of the students in Finland agree that they often think about how the solution to a problem in mathematics can be applied to other interesting questions. In all the Nordic countries, a high proportion of the students agree that they try to understand new concepts by relating them to things they already know. The variation between the Nordic countries on this question is relatively small.

Figure 5 *Elaboration strategies in mathematics: Percentages of the students who 'Strongly agree' or 'Agree'*
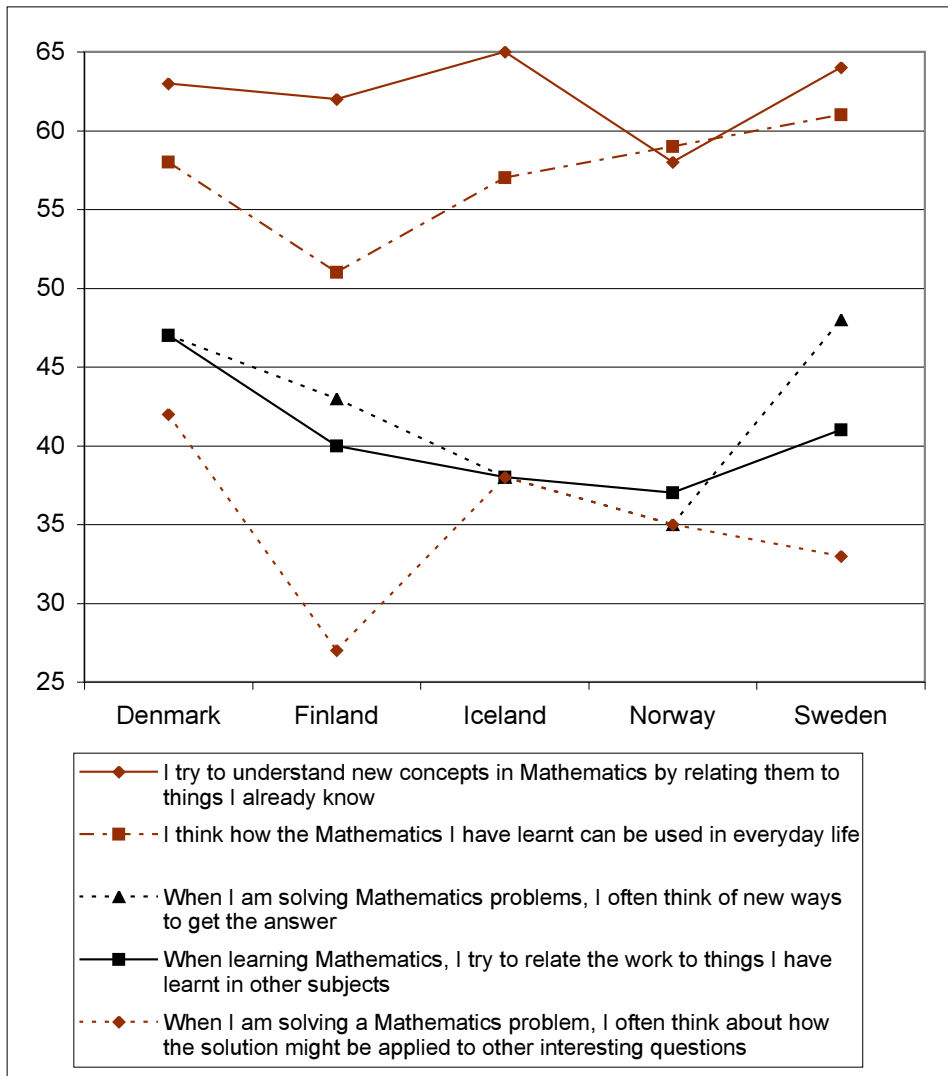
Figure 6 shows large variations between the Nordic countries for some of the items included in the construct 'control strategies'. 75% of students in Iceland agree that they start by working out exactly what they need to learn when they study mathematics. The corresponding result for Sweden is only 40%. Many students in the Nordic countries agree that they work out which are the most important parts to learn when studying for a mathematics test, and that they figure out which

Figure 6 *Control strategies in mathematics: Percentages of the students who 'Strongly agree' or 'Agree'*
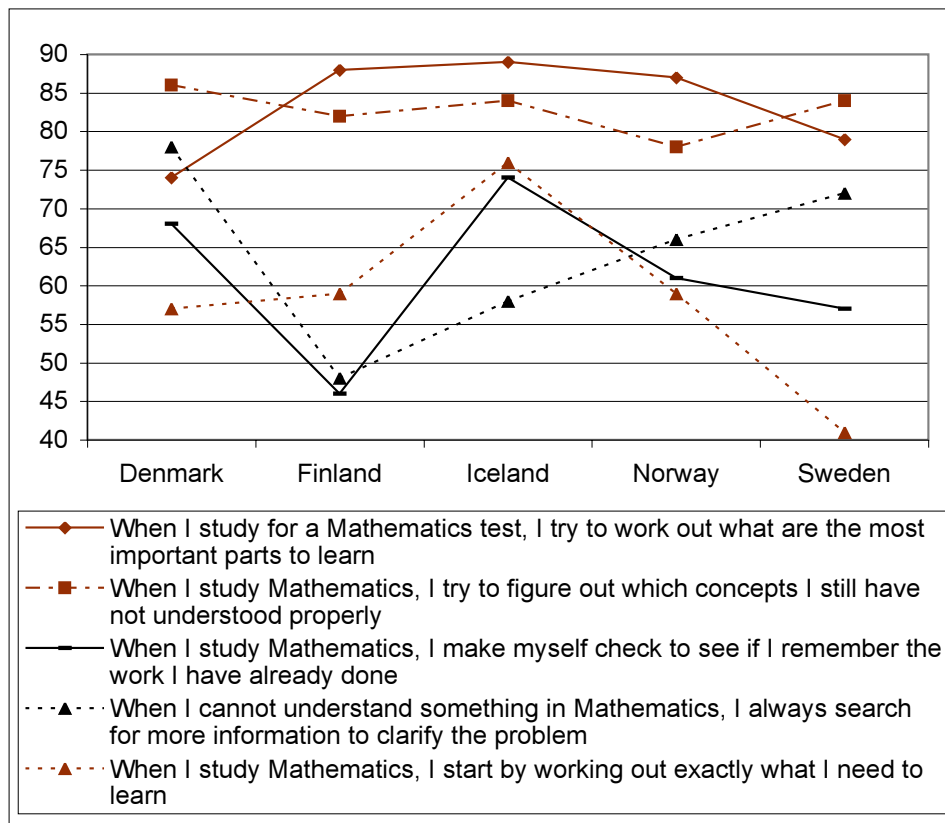
concepts they still have not understood properly when studying mathematics. The variation between the Nordic countries is relatively small on these two questions.

Are the PISA measures directly comparable between countries?

In cross-cultural studies with variables such as values, attitudes and habits, it is often assumed that differences in scores can be compared at face value. However, response styles like acquiescence and extreme response style may affect answers, particularly on rating scales (Artelt et al. 2003; Fischer, 2004; Herk et. al. 2004). It has been argued that cross-cultural questionnaire studies are always challenged by the issue of response biases. Different cultural explanations of response bias have been suggested (Bempechat et al. 2002).

Lie and Turmo (2005) aggregated data from the PISA 2003 student questionnaire at country level and found strong negative correlations between the mean construct values and mean mathematics achievement. These findings contradicted what would have been theoretically expected. They also conducted factor analysis of all the PISA 2003 constructs at country level. The analysis showed that the first component alone could explain 66% of the variance. This was a rather surprising finding, taking into account that the constructs are substantially very different ('sense of belonging', 'elaborations strategies in mathematics', 'instrumental motivation' etc.). However, all the constructs are measured using Likert scales. It was suggested that the first factor in the principal component analysis therefore might be interpreted as the general response tendency on Likert scales in the different countries.

Based on the first component from the principal component analysis, a meta-construct was generated as the linear combination of the constructs. The countries' scores on this meta-construct may be interpreted as a quantitative measure of the general response tendency in each country. The results showed that Tunisia, Brazil and Mexico are the countries with the strongest general tendency to agree to statements, while Japan and Korea had the weakest general agreement tendency.

Similar analysis was also conducted on the data from PISA 2000. The correlation between the meta-construct values in PISA 2000 and PISA 2003 for the countries that participated in both cycles was 0.89. In other words, the agreement tendencies seem to be reasonably consistent between the two PISA cycles. The results showed that Denmark has the strongest general agreement tendency among the Nordic countries, while Finland and Norway have the weakest tendencies.

These findings indicate that direct comparisons of country mean values for the constructs based on Likert scales in PISA, like the learning strategies, should be made with caution. In this sense, the findings are in line with previous research on cross-country comparisons of questionnaire data (Flaskerud, 1988; Heine et al. 2002; Lee et al. 2002). However, in the Nordic countries the estimated correction factors are relatively moderate. This implies that the cultural bias when comparing the results from the Nordic countries, as in this article, is also relatively moderate. However, in countries like Japan, Korea, Brazil and Mexico interpretations of the mean construct values from an international perspective may change significantly if the rather large bias estimates are taken into consideration.

Conclusion

What can we learn from the students' self-reported strategy use in PISA? Most noticeable perhaps is that for students' in the Nordic countries reported use of learning strategies is below the OECD mean. However, country-specific general response tendencies seem to be present in the data from the PISA student questionnaire. The estimated correction factors in the Nordic countries are, however, relatively moderate. Analyses indicate that the student response behaviour is fairly consistent across the Nordic countries, i.e. the *general response tendencies* on Likert scales are rather similar. However, for some of the learning strategies items there are rather large *specific* differences in the students' responses between the Nordic countries.

The results in Finland are perhaps of particular interest. Finland is among the countries with the highest mathematical literacy score in PISA, only outperformed by Hong Kong among all the participating countries. However, the results show that Finnish students' reported use of all the three learning strategies is below the OECD mean, and also below that of students in most of the other Nordic countries. Finland has a particularly low value for control strategies in mathematics, which is rather surprising. Even after correcting for the general response tendency in Finland on Likert items, this finding still holds. The research literature suggests that metacognition is essential for effective learning. It is therefore relevant to ask if there are important aspects of Finnish students' use of strategies in mathematics that are not captured by the measures in PISA. Knain & Turmo (2003) suggest that it is not so much the frequency of use of the strategies that identifies a student who can self-regulate his or her learning, but the fact that the student can flexibly adapt strategies according to the situation. One of the key aims of international studies like PISA is illustrated by the slogan 'learning from others' (Shorrocks-Taylor & Jenkins, 2000). If other countries decide to use Finnish students as examples to learn from, which would be highly relevant based on their high mathematical literacy level, a more elaborated description of their approaches to mathematics learning is needed.

The results for the individual items used to measure learning strategies also show some interesting features, not least among answers to some of the questions used to measure elaboration strategies. Only a quarter of Finnish students agree that they think about how the solution to a problem in mathematics might be applied to other interesting questions. This may reflect a rather abstract and 'pure' approach to mathematics teaching in Finland. On the other hand, more than 60% agree that they try to understand new concepts in mathematics by relating them to things

they already know. Such findings underline the value of studying and discussing the results for learning strategies in PISA at the single item level.

Regarding the within-country correlations between the learning strategies and mathematical literacy, interesting differences between the Nordic countries are found. For example, we have seen that the strongest correlation between memorisation strategies and mathematical literacy is found in Norway, while no relationship is found in Iceland. It is interesting to reflect on how differences in approaches to mathematics teaching in the different Nordic countries might influence this relationship. It has been argued that the time spent on teaching fundamental skills in mathematics is rather low in Norwegian primary and lower secondary schools, and this was also evident in the Norwegian results in TIMSS 2003 (Grønmo et al. 2004). This highlights the importance of the students' own emphasis on rehearsal strategies, as reported in the empirical results in PISA.

Boys report that they use memorisation strategies and elaboration strategies more than girls in all the Nordic countries. What these gender differences mean should definitely be studied in more depth. There may be a gender biased response tendency, with girls tending to answer more modestly than boys. However, no differences in use of control strategies between the genders are found in the Nordic countries. It is interesting to note that the girls report significantly more use of these strategies than boys in 22 of the 30 OECD countries participating in PISA 2003 (OECD, 2004).

The gender difference in the use of memorisation learning strategies varies between the Nordic countries. The difference is especially large in Norway, followed by Denmark and Sweden. In Iceland, the gender difference is moderate. However, the gender difference in the use of elaboration strategies is relatively consistent across the Nordic countries, with boys reporting greater use of elaboration than girls. The practical realities behind these patterns should be studied further.

In summary, our analysis shows that there are a number of questions which remain unanswered regarding the students' self-reports on learning strategies in PISA and how well the scores actually reflect the students' use of learning strategies. Therefore, the empirical results should be interpreted with caution. Validation of the questionnaire instrument using other research methods and approaches, as for example has been done by Samuelstuen (2005), is therefore needed. In Norway, Hopfenbeck will continue this line of research through her PhD work which involves interviewing students about the PISA questionnaire. The goal of the interviews is to collect the students' interpretations of the PISA items and their fundamental thoughts regarding their choice of strategies for solving particular

problems. Individual interviews will be carried out immediately after the students complete the PISA 2006 test.

References:

Artelt, C., Baumert, J., Julius-McElvany, N., & Peschar, J. (2003). *Learners for Life. Student Approaches to Learning. Results from PISA 2000.* Paris: Organization for Economic Co-Operation and Development.

Baumert, J., Fend, H., O'Neil, H.F., & Peschar, J.L. (1998). *Prepared for Life-Long Learning: Frame of Reference for the Measurement of Self-Regulated Learning as a Cross Curricular Competence (CCC) in the PISA Project.* Paris: Organization for Economic Co-operation and Development.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tilmann, K. J., & Weiss, M. (2000). *Self-Regulated Learning as a Cross-Curricular Competence.* Berlin: Max Planck Institut für Bildungsforschung.

Bempechat, J, Jimenez, N.V., & Boulay, B.A. (2002). Cultural-Cognitive Issues in Academic Achievement: New Directions for Cross-National Research. In Porter, A. C., & Gamoran, A. (Eds.): *Methodological Advances in Cross-National Surveys of Educational Achievement.* Washington: National Academy Press. (pp.117-149).

Chamot, A. U. (Ed.). (1999). *The Learning Strategies Handbook.* White Plains, NY: Longman.

Fischer, R. (2004). Standardization to account for cross-cultural response bias. A classification of Score Adjustment Procedures and Review of Research in JCCP. *Journal of Cross-Cultural Psychology*, 35, 263-382.

Flaskerud, J.H. (1988). Is the Likert scale format culturally biased? *Nursing Research*, 37, 185-186.

Grønmo, L. S., Bergem, O.K., Kjærnsli, M., Lie, S., & Turmo, A. (2004) *Hva i all verden har skjedd med realfagene? Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2003.* Oslo: Institutt for lærerutdanning og skoleutvikling, Det utdanningsvitenskapelige fakultet, UiO

Heine, S.J., Lehman, D.R., Peng, K.P., & Greenholz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference group effect. *Journal of Personality and Social Psychology*, 82, 903-918.

Herk, H., Poortinga, Y.H., & Verhallen, T.M.M. (2004). Response Styles in Rating Scales. Evidence of Method Bias in Data From Six EU Countries. *Journal of Cross-Cultural Psychology*, 35, 346-360.

Knain, E., & Turmo, A. (2003). Self-regulated learning. In Lie, S., Linnakylä, P., & Roe, A. (Eds.) *Northern Lights on PISA. Unity and Diversity in the Nordic countries in PISA 2000.* Oslo: Department of Teacher Education and School Development, University of Oslo.

Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, W. E. (2002). Cultural Differences in Responses to a Likert Scale. *Research in Nursing & Health*, 25, 295-306.

Lie, S., & Turmo, A. (2005). *Cross-country comparability of students' self-reports.* Paper to the PISA Technical Advisory Group Meeting, TAG (0505)11, Paris, France, May 9 -11 2005.

OECD (2004). *PISA 2003: Learning for Tomorrow's world. First results from PISA 2003.* Paris: OECD Publication.

OECD (2005). *PISA 2003: Technical Report.* Paris: OECD Publications.

Ramsden, P (1992). *Learning to teach in higher education.* London: Routlage Falmer

Samuelstuen, M. (2005) *Kognitiv og metakognitiv strategibruk med særlig henblikk på tekstlæring.* Trondheim: Doktoravhandling for graden doctor rerum politicarum, Fakultet for samfunnsvitenskap og teknologiledelse, NTNU.

Shorrocks-Taylor, D., & Jenkins, E. W. (Eds.) (2000) *Learning from Others.* Dordrecht: Kluwer Academic Publishers.

Turmo, A. (2005). The relationship between the use of learning strategies and socioeconomic background in 15-year olds. *Journal of Nordic Educational Research*, 25, 155 -168.

Weinstein, C. E. (1988). Assessing Learning Strategies: The Design and Development of the LASSI. In Weinstein C. E., Goetz, E., T., & Alexander, P. A. (Eds.) *Learning and Study Strategies Issues in Assessment, Instruction and Evaluation.* San Diego: Academic Press (pp. 26-39)

Zimmerman, B. J. (1990). Self-regulated Learning and Academic Achievement: An Overview. *Educational Psychologist,* 25, 3 -17.

Chapter 9

Nordic Minority Students' Literacy Achievement and Home Background

Astrid Roe and Rita Hvistendahl

Abstract

This chapter presents the results of Nordic minority students in mathematics, reading and science in PISA 2003.[1] On average minority students achieve lower scores than majority students in all three domains of the assessment, and they also have parents with a lower socio-economic status than majority students. However, there are some interesting differences between countries and between groups of students. Minority students have been divided into two groups: those born in the country of assessment and those born in another country. The test scores of majority students and the two groups of minority students for the three domains of assessment show different patterns in each of the Nordic countries. Socio-economic status and students' attitudes towards school also vary between countries and between groups of students.

Nordic abstract

Blant de elevene som deltok i PISA 2003, var det også elever med et annet morsmål enn det offisielle språket i de enkelte landene. Mange av disse elevene snakker morsmålet sitt hjemme, mens de må lese, snakke og forstå det offisielle språket i landet - både på skolen og i samfunnet for øvrig. I de nordiske landene varierer prosentandelen elever med minoritetsspråklig bakgrunn, som deltok i PISA 2003, fra én prosent av elevene på Island til elleve prosent av elevene i Sverige. Gjennomsnittlig skårer minoritetselevene lavere enn majoritetselevene på alle de tre fagområdene i undersøkelsen, men det er store variasjoner mellom elever som er født i landet, og de som er født i et annet land, og minoritetselevenes foreldre har gjennomsnittlig lavere sosioøkonomisk status enn foreldrene til majoritetselevene. I den første delen av denne artikkelen presenterer vi en sammenlikning av prestasjonene i matematikk, lesing og naturfag mellom nordiske

---

1. Minority students are defined here as students with both parents born outside the country of assessment.

majoritetselever, minoritetselever som er født i landet, og minoritetselever som ikke er født i landet. Vi presenterer deretter resultater fra elevspørreskjemaet, basert på spørsmål om hjemmebakgrunn, læringsstrategier og motivasjon. Vi ser også på sammenhengen mellom noen av bakgrunnsvariablene og prestasjoner i de tre fagområdene.

Introduction

Among the students who participated in PISA 2003 were students whose first language is different from the language of assessment. Many of them speak their first language at home, whereas they have to read, speak and understand the official national language at school and in society. Within the Nordic countries the percentage of students taking part in PISA 2003 who come from an immigrant background varies from 1 per cent of the student sample in Iceland to 11 per cent in Sweden. In the first part of this chapter we will present a comparison of achievement in mathematical, reading and scientific literacy between majority students, immigrant students born in the country of assessment and immigrant students born elsewhere for the Nordic countries. We will also present results from the PISA student questionnaire about family background, learning strategies and motivation at school. Finally, we will present correlations between some background variables and literacy achievement.

Immigration to the Nordic countries

A study of the consequences of immigration on the welfare policies of the Nordic countries (Brochmann and Hagelund 2005) gives an overview of immigration to these countries in the period after 1945. The study points out that the history of immigration in Denmark is in many ways typical of the overall immigration history of the Nordic countries. In the 1960s and 70s Denmark experienced immigration of labourers from Turkey, Pakistan and Yugoslavia. In the 1970s immigration was restricted as a result of the oil crisis and subsequent unemployment, but continued from countries in the third world through family reunion and the acceptance of refugees. Today almost 6 % of the population in Denmark originates from countries outside the EU, the Nordic countries and North America. Politically Denmark has reacted to immigration somewhat differently from other Nordic countries by decentralizing the settlement of immigrants and restricting family reunion from third world countries.

Among the Nordic countries Sweden has the largest immigrant population, 15 % of the total population. Modern immigration to Sweden started early in the post-war period, and since 1945 Sweden has accepted more than 1 million immigrants. The period from the Second World War to the mid-1970s was dominated by labour immigration, while most immigrants in subsequent years,

particularly from the 1990s, have been refugees, asylum seekers and relatives joining their families, many of them from East Europe and other countries outside the OECD. By the late 1960s Sweden had already adopted a policy of integrating immigrants into Swedish society, which emphasised equal opportunities, justice, solidarity and cooperation, and within the Nordic countries Sweden has pioneered the integration of immigrants. In spite of this effort, Sweden has no more success at integrating newcomers today than other Nordic countries.

Immigration to Norway from remote countries such as Morocco, Turkey and Pakistan started later at the end of the 1960s, and was more limited than in Sweden. Today 7.6 % of the total population in Norway has an immigrant background. The nationality groups within the immigrant population of Norway are the same as in Denmark and Sweden, except for a large concentration of immigrants from Pakistan. Norwegian immigration policy was highly influenced by Swedish policy until the end of the 1980s, but subsequently the two countries have adopted different policies.

Population movement in Finland has mainly involved emigration and to a small extent labour immigration, and the country has received few refugees. Many of the immigrants to Finland have come through marriage to Finnish citizens, and not as a consequence of an active immigration or refugee policy. Only 2 % of the total population has an immigrant background. Many immigrants are of Finnish origin, and more than a third come from Russia and Estonia.

Until 1995 Iceland had a negative immigration rate. Since then immigration has fluctuated. The immigrant population is today approximately 10 000, with immigrants from Poland forming the largest nationality group followed by immigrants from Denmark, Germany and the Philippines. Very few immigrants are refugees. Modern immigration to Iceland is mainly labour immigration.

The different patterns of immigration to the Nordic countries and differences in integration policies are likely to influence the minority students' family background by affecting resources at home and the parents' socio-economic status. The results of the PISA study 2003 must therefore be interpreted in the light of similarities and differences in immigration to the Nordic countries, as well as in the countries' policies on integration.

Minority students in PISA 2003

Two different criteria were used to categorise students with an immigrant background who participated in PISA 2003: (A) parents' and students' countries of birth and (B) the language spoken at home most of the time. In the following presentation we will use the categorisation based on the first criterion. However, many students from group A will also be found in group B. Students were divided into the following three groups (OECD, 2001, pp. 293-294):

Majority students: students born in the country of assessment with at least one of their parents born in the same country

Minority students born in the country of assessment with foreign-born parents

Minority students not born in the country of assessment with foreign-born parents

In the following section, the term *minority students* will be used about students with foreign-born parents, regardless of the student's country of birth. A distinction between minority students born in the country of assessment and those born in foreign countries will be made in the presentation of the results in reading, mathematical and scientific literacy in Denmark, Norway and Sweden. Otherwise the two groups of students will be referred to as minority students and majority students.

A total of 1127 Nordic minority students participated in PISA 2003. Table 1 shows the distribution of minority students in each Nordic country (the percentage of the total number of students is given in parenthesis). It also shows the distribution of minority students born in the country of assessment and those born elsewhere.

Table 1 *Total number of minority students in each of the Nordic countries, also shown as a percentage of the total number of students in parenthesis*

	Denmark	Finland	Iceland	Norway	Sweden
Total number of minority students	258 (6.1%)	99 (1.7%)	33 (1%)	227 (5.6%)	510 (11%)
Students born in the country	137	3	5	95	245
Students not born in the country	121	96	28	132	265

The table shows that there is a higher percentage of minority students in Denmark, Norway and Sweden than in Finland and Iceland, and that Sweden has the largest proportion of minority students among the 15-year-olds who participated in PISA. The table also shows that most of the Finnish and Icelandic minority students were born in another country, whereas the two groups are more evenly distributed in the Scandinavian countries, but with more minority students born abroad in Norway than in Sweden and Denmark.

Minority students' achievements

Figure 1 shows the average results in reading, mathematical and scientific literacy for Nordic majority students, minority students *born* in the country of assessment and minority students *born* in another country[2]. The results for all five countries

are combined into one group in the figure. The gap between majority students and minority students born in the country of assessment is significant for all three domains. Minority students who are born in another country generally perform about 25 points below minority students who are born in the country. If we look at the same results for each country separately, a more nuanced picture appears.

Figure 1 *Figure 1 Average score points in mathematics, reading and science for three groups of students in all Nordic countries*
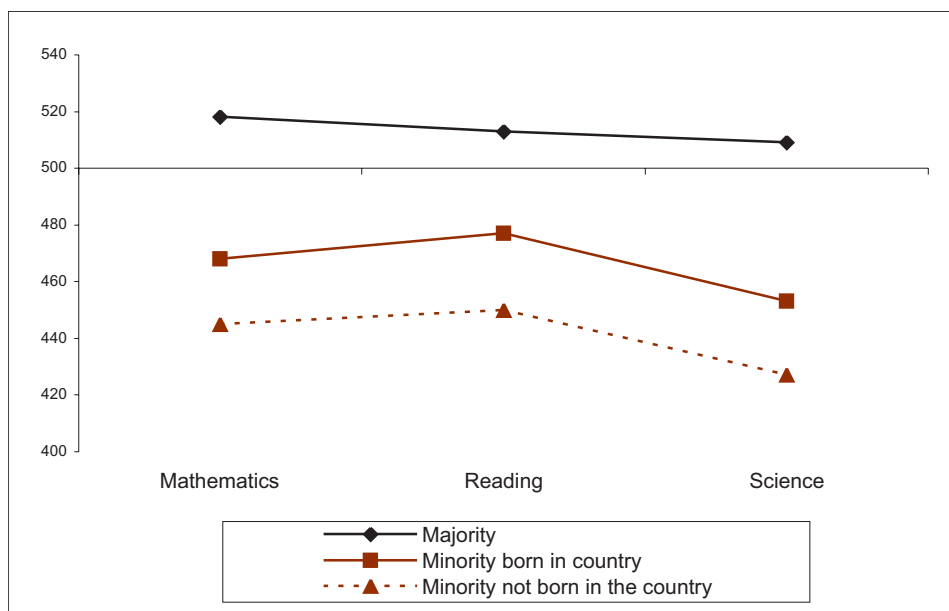
In the Finnish and Icelandic PISA sample there are very few minority students who are born in the country of assessment, so results for these countries will not be presented in figures 2, 3 and 4. Figure 2 shows the mathematics results for the three groups of students in Denmark, Norway and Sweden. In Denmark there is a large achievement gap between majority students and minority students in general, regardless of their country of birth. In Norway the relative differences between the three groups of students are quite similar to the total Nordic results in figure 1. Swedish results are characterised by a moderate achievement gap between majority students and minority students born in the country, whereas Swedish minority students born in another country perform quite poorly compared to those born in Sweden.

---

2. Achievement scores in figures 2 – 4 are based on warm estimates

Figure 2 *Average score points in mathematics for three groups of students in Denmark, Norway and Sweden*
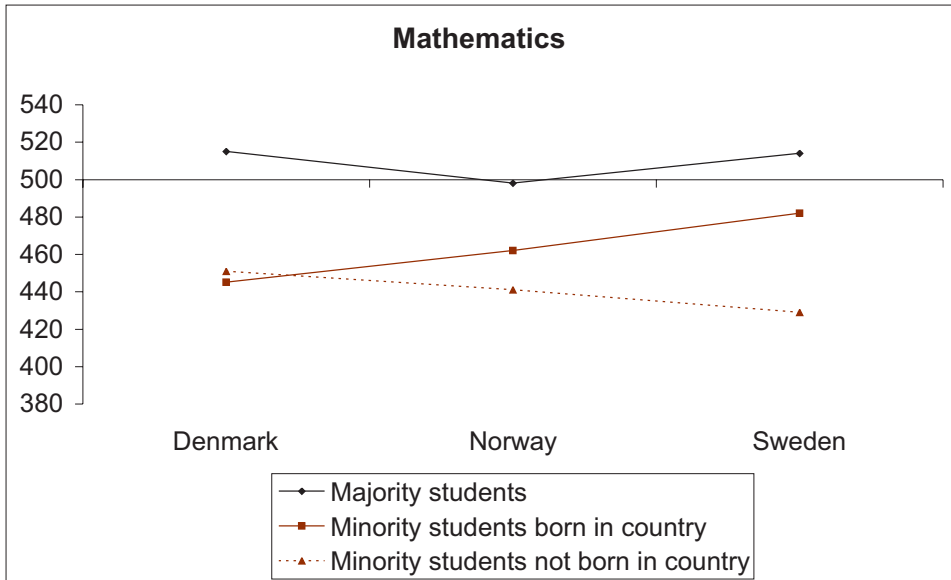
Figure 3 shows that there is a significant gap between majority students and both groups of minority students in Denmark and Norway. As far as the reading proficiency of Norwegian minority students is concerned, being born in Norway seems to give a small advantage compared to being born in another country. In Denmark, the situation seems to be the opposite, which is somewhat surprising. However, given that the standard errors for the minority groups are of the order of 10 score points, one should bear in mind that these differences can partly be explained by sampling errors. In Sweden, minority students born in the country perform very well, but there is a significant gap between the two minority groups.

The Norwegian and Swedish results in figure 4 show that science seems to be an even greater challenge than reading and mathematics for minority students who are not born in the country. The Danish results show the same tendency as in figure 3, namely that minority students born in Denmark perform worse than those who are born abroad, although the difference is not statistically significant.

The fact that Swedish minority students born in Sweden seem to perform so well in reading and scientific literacy indicates a relation between achievement and language proficiency. The reading and scientific literacy tasks in PISA both require better reading competencies and language proficiency than the mathematical

Figure 3 *Average score points in reading for three groups of students in Denmark, Norway and Sweden*
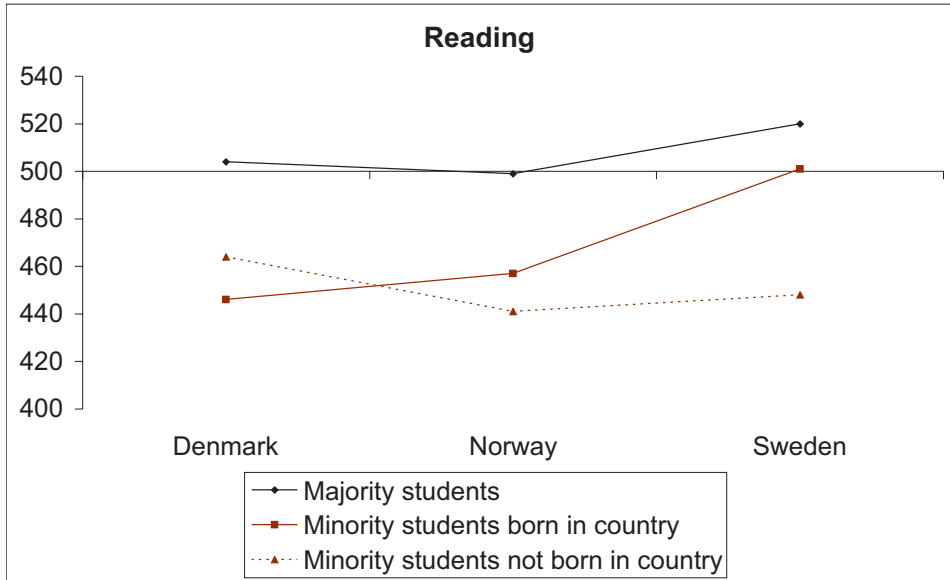
Figure 4 *Average score points in science for three groups of students in Denmark, Norway and Sweden.*
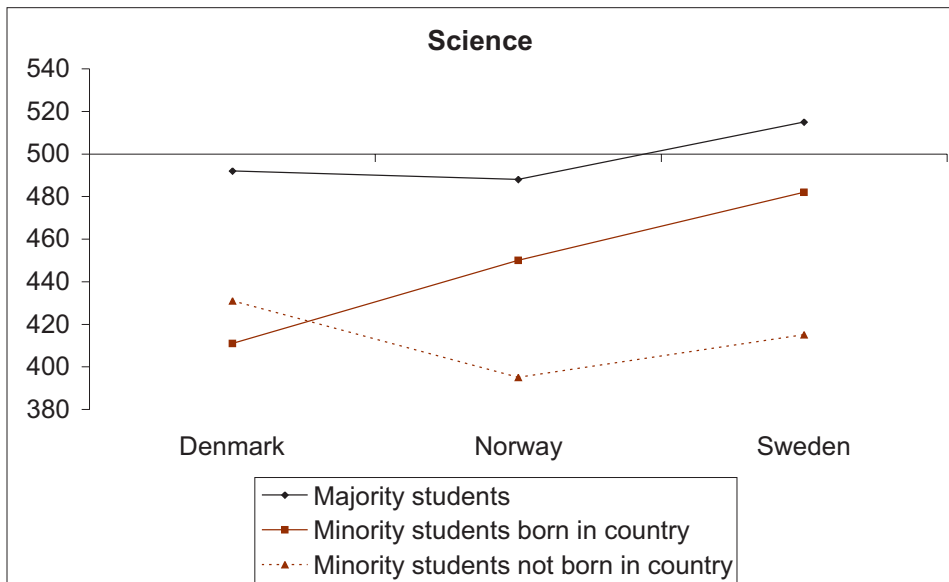
literacy tasks. The science tasks are particularly demanding, because they require a combination of language and content knowledge. All texts and tasks in PISA are presented in the national language, which for most minority students is their second language. Therefore most minority students will have their second language reading competencies assessed in PISA, whereas most majority students will have their first language reading competencies assessed. A study conducted by Kulbrandstad (1996, p. 450) states that reading in a second language is slower and generally involves a lower degree of understanding than reading in a first language. The good Swedish results for minority students born in the country may either indicate that the Swedish school system has developed better methods as regards systematic instructional practices for minority students than the other two countries, or that minority students born in Sweden have a better command of the national language when they start school than minority students born in Norway and Denmark.

Minority students' family background

A lot of data about students' family background and attitudes towards school were collected through a questionnaire administered to students. This offers the possibilities of comparing Nordic minority students' home backgrounds and attitudes towards school, and of studying the relation between these background variables and achievement. In the following section, differences between minority and majority students regarding some of these variables are reviewed, and the correlation between the variables and the students' literacy achievements is analysed. Correlations between achievement and all the background variables revealed that some variables were more significantly correlated with achievement than others. The following variables showed the most significant correlation with achievement in the Nordic countries: *computer facilities at home*, *cultural possessions at home*, *educational resources at home*, *number of books at home*, and *parents' highest socio-economic status*.

Resources at home

Computer facilities at home, *cultural possessions at home and educational resources at home* are all composite variables with evaluation of each variable consisting of a number of single questions. The 0 level in figure 5 represents the OECD average and the standard deviation is 1. The figure shows that the level of Nordic majority students' cultural possessions at home is moderately above or near the OECD average, except for the Icelandic students, who seem to come from very well equipped homes as far as cultural possessions are concerned. Nordic majority students also seem to have more educational resources at home than the OECD average, with the exception of Danish students who are significantly below this

average. Nordic majority students are generally well off in terms of computer facilities at home compared to the OECD average. The computer facilities of Finnish minority students seem to be nearly as good as the OECD average, and in the other four countries minority students have computer facilities even better than the OECD average. The case is different for cultural possessions and educational resources. With the exception of cultural possessions in Iceland and educational resources in Norway, Nordic minority students are below the OECD average.

Figure 5 *Possessions at home among majority and minority students in each Nordic country. (OECD mean = 0, one standard deviation = 1)*
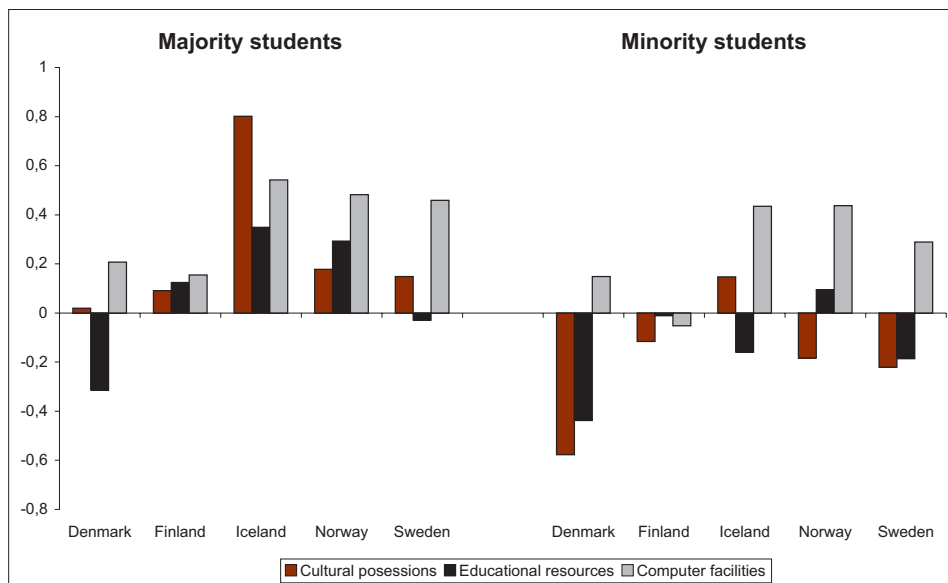
Among the variables analysed, *cultural possessions* was the most significantly different between majority and minority students in all Nordic countries. The composite variable *cultural possessions* included single questions about whether the student has classical literature (e.g. Shakespeare in the English version), or collections of poems and visual art (e.g. paintings) at home. The cultural possessions mentioned in the questions are typical status symbols related to Western culture. The PISA study mainly defined "classical" culture in terms of Western elite culture. This definition did not include cultural possessions or activities highly valued among minority students and their families.

Number of books at home is a variable that is frequently used to measure the socio-economic status of families. In the earlier Third Mathematics and Science Study (TIMSS) this variable showed significant correlation with achievement in science

Figure 6 *Number of books at home among majority and minority students in each Nordic country*
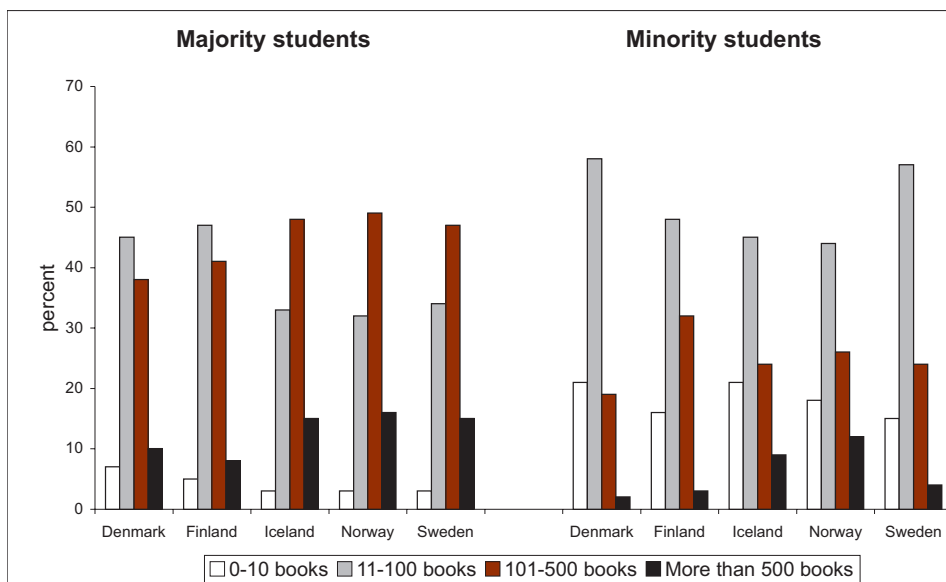
and mathematics (i.e. Lie *et al.*, 1997, p.182). The PISA 2000 data showed a positive correlation (0.32) between number of books at home and achievement for majority students. The PISA 2003 data revealed that minority students in all the Nordic countries, have fewer books at home than majority students, as shown in figure 6. This is not surprising given the migrant situation and the socio-economic status of the minority students' families.

Parents' occupational status

Education and occupation are related to other home background factors: In PISA 2003 the highest family occupational index is derived from the parental occupations given in the student questionnaire. The following questions are asked about both parents' professions:

- *What is your mother's main job? (e.g. school teacher, nurse, sales manager). If she is not working now, please tell us her last main job. Please write in the job title.*
- *What does your mother do in her main job? (e.g. teaches high school students, cares for patients, manages a sales team).*
- *If she is not working now, please tell us her last main job. Please use a sentence to describe the kind of work she does or did in that job.*

The classification of the parents' occupations uses a system designed for classification of occupations across countries (Ganzeboom & Treiman 1996). The

index gives a metric for the sum of education level and income level for each occupational category. This has been quantified based on international empirical data. The scale ranges from 0 to 90, and the level of occupational status increases with increasing values.

Figure 7 shows that the parents of Nordic students' in general have higher occupational status than the OECD average. The figure also shows a difference in the occupational status of the parents between majority and minority students in Scandinavian countries. No such difference is seen between majority and minority students in Finland and Iceland and this may be explained by the differences in immigration to Denmark, Sweden and Norway on the one hand and Finland and Iceland on the other. Immigration from remote countries in the third world to the three Scandinavian countries has been much greater than to Finland and Iceland, and the Scandinavian countries have received significantly more refugees and asylum seekers than the two other Nordic countries.

Figure 7 *Parents' highest occupational status among majority and minority students in each Nordic country, compared to the OECD mean*
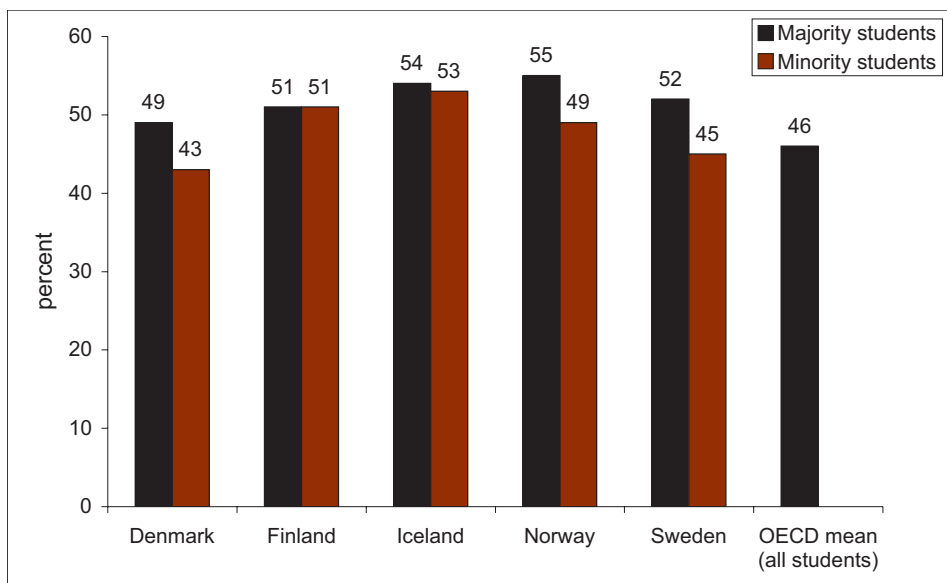
Correlations between home background and achievement

Correlations between variables connected to family background and achievement in reading, mathematical and scientific literacy may reveal to what extent family background is related to minority students' literacy achievements. The variables did not show a significant correlation with scores in Iceland, probably due to the low

number of minority students in this country. Only correlations that are significant at the 0.01 level are shown here. Table 2 shows that in most cases the variables correlated statistically significantly with the scores, with the exception for all three domains of *computer facilities* in Norway and *parents' highest occupational status* in Denmark. For science in Norway there was no significant correlation with *educational resources at home*. The correlations are generally moderate and the differences between countries are difficult to explain. A tentative explanation for the lack of significant correlation between computer facilities at home and achievement in Norway may be that computers are found in most homes, regardless of educational or cultural background. However, it is difficult to explain why this is not also true in Denmark and Sweden. It is even more difficult to explain why parents' highest socio-economic index shows no significant correlation with scores for minority students in Denmark. The correlations are generally quite similar across the three domains within each country. The conclusion must be that family background is important for achievement although the results still leave many questions unanswered. Correlation coefficients for majority students are also shown in the table, but will not be discussed in this chapter.

Table 2 *Correlation coefficients between home background and literacy achievement for minority students and majority students in the Nordic countries*

	Denmark		Finland		Iceland		Norway		Sweden	
Computer facilities at home	Majority	Minority	Majority	Minority	Majority	Minority	Majority	Minority	Majority	Minority
Mathematics	0,19	0,28	0,15	-	0,11	-	0,12	-	0,14	0,24
Reading	0,16	0,19	0,11	-	0,51	-	0,09	-	0,13	0,27
Science	0,18	0,24	0,14	-	0,13	-	0,1	-	0,12	0,24
Cultural possessions at home										
Mathematics	0,31	0,17	0,21	-	0,19	-	0,32	0,23	0,3	0,36
Reading	0,36	0,18	0,25	0,28	0,17	-	0,31	0,24	0,31	0,36
Science	0,35	0,19	0,28	0,25	0,21	-	0,28	0,23	0,28	0,34
Educational resources at home										
Mathematics	0,26	0,21	0,18	-	0,14	-	0,24	0,22	0,25	0,32
Reading	0,3	0,18	0,22	-	0,11	-	0,25	0,21	0,25	0,29
Science	0,26	0,2	0,21	-	0,16	-	0,25	-	0,23	0,32
Number of books at home										
Mathematics	0,31	0,25	0,28	0,3	0,26	-	0,31	0,33	0,34	0,32
Reading	0,3	0,21	0,29	0,29	0,21	-	0,3	0,29	0,34	0,35
Science	0,31	0,27	0,28	0,32	0,26	-	0,29	0,36	0,32	0,34
Parents' highest socio occupational status										
Mathematics	0,29	-	0,26	0,31	0,17	-	0,3	0,31	0,3	0,22
Reading	0,26	-	0,22	-	0,11	-	0,26	0,26	0,26	0,23
Science	0,29	-	0,22	-	0,13	-	0,3	0,35	0,27	0,23
The correlation coeficcients shown in the table are at the 0.01 level.										

School-related factors

An important aim of schooling is to develop competencies that go beyond the school subjects, like *learning strategies, motivation, self-concept* and *learning style* (Lie et al., 2001). *Learning strategies* included *memorisation, elaboration* and *control strategies.*

Figure 8 *School related factors among all Nordic minority and majority students. (The OECD mean = 0. One standard deviation = 1)*
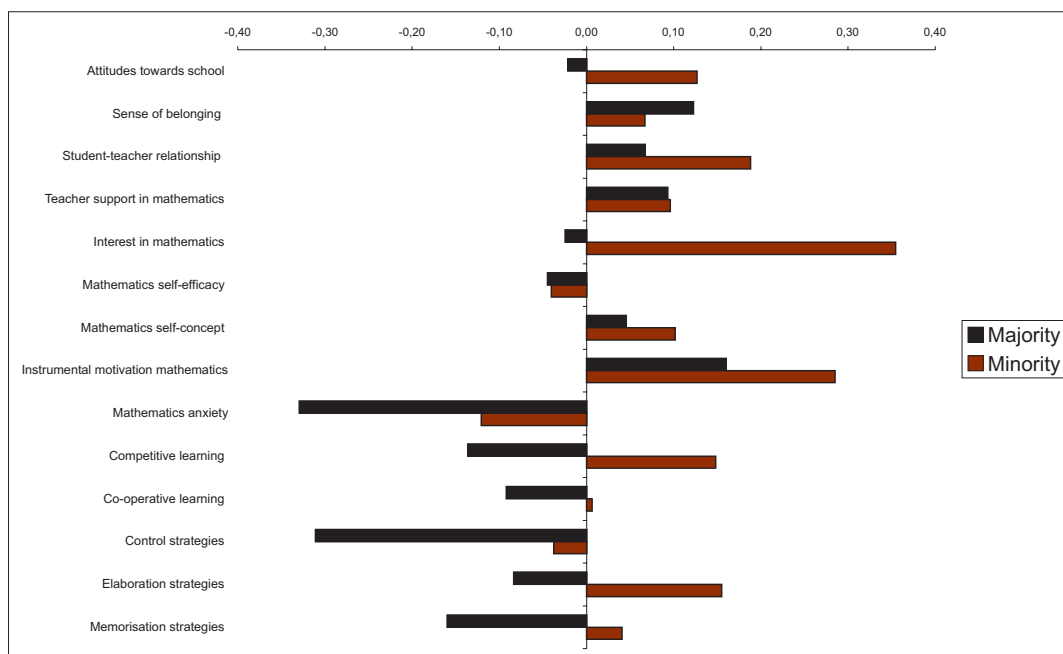
With a couple of exceptions, there are significant differences in these competencies between majority and minority students, as shown in figure 8. The results for students from all the Nordic countries are combined in this figure because, although there are some differences between countries, the differences between majority and minority students in each country go in the same direction. Minority students in the Nordic countries tend to be more motivated, interested and have more positive attitudes towards learning mathematics at school, and they seem to have a stronger belief in the importance of mathematics in education than majority students. They are more confident in mathematics than majority students and they report a better relationship with their teachers. Minority students also seem to be more competitive and co-operative than majority students, and they report more frequent use of learning strategies. Only one variable presents more positively for

majority students; *sense of belonging at school.* Minority students generally report to a stronger degree feeling lonely, awkward and like an outsider. However, they still have a stronger sense of belonging than the OECD average.

Summary and discussion

The average difference in literacy achievement between minority and majority students varies between the Nordic countries. When comparing the achievement gap between the two groups of students in Denmark, Norway and Sweden, we find Swedish minority students born in the country perform best compared to majority students, particularly in reading literacy. The gap between majority students and minority students born in the country of assessment is generally much bigger in Denmark and Norway, with one exception: In science the difference is about the same between the two groups in Sweden and Norway. The results indicate that there are differences in the way schools have mastered the challenge of teaching minority students the official national language in the three countries. The relatively good results of Swedish minority students born in Sweden may also be partly explained by the Swedish immigrant integration policy developed in the early 1960s which emphasises equal opportunity, justice, solidarity and cooperation. Minority students born outside the country of assessment perform at almost the same level in Norway and in Sweden, scoring approximately 450 points in reading literacy. However, the number of foreign-born minority students is proportionately larger in Norway than in Sweden. In PISA 2000 we found that the length of time the minority students' had lived in Norway was of great importance for their literacy achievements. In particular, minority students who started attending school in Norway in the 5th grade or later had lower results. Thus, Norwegian schools face a double challenge of tailoring the literacy education to the needs of both minority students born in Norway and minority students who start school in another country and continue their education in Norway (Hvistendahl & Roe 2004). This challenge is also faced in the other Scandinavian countries.

Students with a different first language from the test language and with foreign-born parents may face severe challenges at school. It is therefore important to be aware of the strengths and weaknesses of this group of students within the school system. The PISA student questionnaire contained questions designed to map socio-economic background, reading habits, and several school-related factors such as motivation and sense of belonging. The reports of Nordic minority students on these questions give a complex picture of their motivation and achievement at school. The fact that the language used at school is not their first language may explain why minority students evaluate their effort and perseverance as high. Minority students also seem more motivated regarding school-related activities

than majority students, which might compensate to an extent for language and cultural disadvantages.

The achievements of the minority students in the Nordic countries reflect to some degree similarities and differences in the pattern of immigration to these countries as well as in their integration policies. It seems that minority students born in Sweden benefit from the country's experience of immigration over many years and its long tradition of forming integration policies that aim to create equal opportunities for the immigrant population. The home background of the minority students reflects current differences in immigration to Finland and Iceland on the one hand and the three Scandinavian countries on the other. The distribution of computer facilities and educational resources shows that minority students and their families benefit to a certain extent from welfare policies and the positive economic situation in most of the Nordic countries, whether they are born in one of these countries or not. The great challenge for all five Nordic countries is to improve the literacy achievements of minority students, especially those who are not born in the country.

References

Brochmann, Grete og Anniken Hagelund, 2005. *Innvandringens velferdspolitiske konsekvenser. Nordisk kunnskapsstatus. TemaNord 2005:506.* København: Nordisk Ministerråd.

Ganzebom, H.B.G. & Treiman, D.J. (1996). Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations*, 25, p. 201-239.*

Hvistendahl, R & Roe, A. 2004. The Literacy Achievement of Norwegian Minority Students. *Scandianvian Journal of Educational Research*, 48 (3) P. 307-324.

Kulbrandstad, L. I. (1996). *Lesing på et andrespråk. En studie av fire innvandrerungdommers lesing av læreboktekster på norsk* [Reading in a second language. A study of four immigrant students' reading of textbook texts in Norwegian; in Norwegian]. Oslo: HF-fak., Universitetet i Oslo.

Lie, S., Kjærnsli, M., Roe, A. & Turmo, A. (2001) *Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i et internasjonalt perspektiv* [Well prepared for the future? Norwegian 15-year-olds' competence in reading, science and mathematics in an international perspective; in Norwegian]. Oslo: Universitetet i Oslo, Institutt for lærerutdanning og skoleutvikling.

Lie, S., Kjærnsli, M. & Brekke, G (1997). *Hva i all verden skjer i realfagene? Internajonalt lys på trettenåringers kunnskaper, holdninger og undervisning i norsk skole*. University of Oslo.

OECD (2001) *Knowledge and Skills for Life. First results from PISA 2000*. Paris: OECD.

How Can Reading Abilities Explain Differences in Maths Performances?

Astrid Roe and Karin Taube

Abstract

This chapter will focus on the influence of reading ability on students' achievements in mathematics in PISA 2003. On the basis of the results in reading and mathematics from 4595 students who participated in PISA 2003 in Norway and Sweden the following questions will be raised and discussed: How do reading abilities influence students' performance in mathematics? What characterises mathematics items that show a high correlation with the overall reading score?

Nordic abstract

Dette kapitlet fokuserer på hvordan leseferdigheter kan påvirke elevers prestasjoner i matematikk i PISA 2003. På bakgrunn av resultatene i lesing og matematikk for 4595 elever som deltok i PISA 2003 i Norge og Sverige, blir følgende spørsmål reist og drøftet: Hvordan påvirker leseferdighetene elevenes prestasjoner i matematikk? Hva er det som karakteriserer de matematikkoppgavene som viser en høy korrelasjon med lesing?

Introduction

In PISA mathematical literacy is defined as "an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen" (OECD, 2003). Thus, the PISA mathematical literacy domain studies the students' abilities to "analyse, reason and communicate ideas and interpret mathematical problems" (p. 15). One can easily understand that in this context not only is the students' mathematical knowledge important but also their ability to read and write.

According to Niss and Højgaard Jensen (2002) the ability to interpret and understand mathematical texts is included in mathematical knowledge. Fuentes (1998) and Cowen (1991) both emphasize the need for students to learn how to read and understand mathematical texts. Möllehed (2001) found that among students in grades 4 to 9 attempting to solve mathematical problems the most common mistakes were related to text comprehension.

The main differences between everyday language and mathematical language are mathematical symbols and the degree of precision (Jakobsson-Åhl, 1999). Mathematical thinking is expressed in words, sentences and symbols where the language often is very precise and where all short words have to be interpreted correctly in order not to distort the meaning (El-Naggar, 1996). In a study by Österholm (2004) 106 participants used three different texts: one history text about the Russian revolution and two mathematical texts about group theory (one with and one without mathematical symbols in the presentation). After reading the texts the participants answered questions about the content. The findings showed that those who read the mathematical text without symbols had better results than those who read the mathematical text with symbols. Österholm concluded: "Thus, the reading of mathematical texts with symbols is a rather special activity and there might be a need for learning how to read such texts" (p.vii). A comparison between the responses of participants reading the historical text and the mathematical text without symbols, on the other hand, showed many similarities. Otterburn and Nicholson (1976) found that many words frequently used in the students' maths books were difficult to understand for a great number of students. According to Chinn and Ashcroft (1998) students who "lose" small words, transpose words in sentences or have a low ability to decode words correctly and efficiently find it hard to understand the logic in texts. They need the help of a teacher to practise interpreting and understanding mathematical texts. Shuard and Rothery (1988) found that many mathematical concepts that teachers and authors expect students to understand simply do not exist in their vocabulary. Students with reading and writing disabilities were sometimes found to copy procedures and to work in a mechanical way (Miles, 1992). According to Kibel (1992) one of the reasons behind the difficulties many students have understanding mathematics is that the symbols are introduced too early in their education.

Besides problems with mathematical concepts and words, sentence structure may cause difficulties (Shuard & Rothery, 1988). Sentences that are structured in the way that students talk are the easiest to understand, while sentences written in the passive tense and subordinate clauses are more complex and difficult to comprehend (Perera, 1980). Graphic material such as pictures, diagrams and tables is an important, frequently used component of mathematics text books. Some

categories of common difficulties in interpreting graphic illustrations are presented in a study by Åberg-Bengtsson (1994).

Hubbard (1992, p.81) criticises the normal style of writing for mathematical text: "the need for the text to be absolutely mathematically correct and complete so that it cannot be criticized by mathematical colleagues (…) results in texts written for mathematicians, not students". Miles (1992) emphasises the importance of the teacher's understanding of how linguistic problems influence the students' learning in mathematics. This understanding is crucial for the students' self-confidence and motivation to learn mathematics. To summarise, earlier research has found considerable evidence that students experience difficulties with words, symbols, sentence structure and graphic material when trying to solve problems defined as mathematical problems.

This chapter will focus on the influence of reading ability on students' achievements in mathematics in PISA 2003, based on the results in reading and mathematics assessments in Norway and Sweden. Separate analyses of the results of Norwegian and Swedish students showed that the differences between the two groups are very small and generally of no statistical significance. Therefore they will be treated as one group in this study. We will not focus on the differences between the Swedish and the Norwegian languages, although we have noticed some translational differences between the Swedish and the Norwegian versions, which may have influenced reading comprehension positively or negatively in one country or the other. However, 'good' and 'bad' translations seem to be evenly distributed between the two countries. In the linguistic analyses of text presented later the English source versions of the PISA material will be used.

Correlations between reading and mathematics

A comparison between the overall scores in reading and mathematics in PISA 2003 showed a correlation coefficient of 0.57[1], which indicates that there is a close relationship between students' reading abilities and their performance in mathematics. This may be explained by the fact that reading and mathematics are both parts of the general concept of "literacy", which is dominant in each assessment area in PISA. In PISA, literacy has a much broader meaning than the ability to read and to write. A literate person has a range of competencies, and the framework of each assessment area in PISA presupposes that fifteen-year-olds have

---

1. Correlations are based on Warm Estimates in Mathematics and Reading, i.e. they include only the 4595 Norwegian and Swedish students who had booklets with Reading and Mathematics in PISA 2003.

a solid foundation of knowledge (OECD 2004). The high correlation may also be explained by the fact that the test items in all assessment areas in PISA are organised in units based on written passages or graphics. Students have to read, understand and interpret both the written passage and the question before they can actually demonstrate their knowledge. Thus, reading plays an important part in all content areas in PISA. It is also important to bear in mind that reading in PISA is defined in terms of the individual's ability to use written text to achieve his or her purpose – which means the capacity to retrieve information, understand, interpret and reflect on a text (OECD 2003). However, this implies that these higher order functions must be based on the ability to decode texts and to understand their crucial concepts. Students who fail to solve mathematical problems might do so because they are unable to do one, some or all of the following: correctly decode words, understand their exact meaning in a mathematical context, reflect on the mathematical problem, actually solve the problem and present the solution to the problem in written words so that others can understand it. Mathematics teachers have to be aware of the fact that when students face written mathematical tasks in daily life it is not only purely mathematical difficulties that present obstacles.

The influence of reading on mathematics performance

In PISA 2003 there are 53 mathematics units with a total of 84 items. The mathematics tasks in PISA are text-based, in the sense that students have to read and understand a certain amount of text to be able to solve the mathematical problem in each item. We assume that some items would be more text-based and/or more difficult to read and understand than others. A possible degree of reading difficulty in the mathematics tasks can be investigated by correlating the test scores for each mathematics item with the overall test score in reading. An item-by-item correlation with reading scores showed positive correlation coefficients between 0.1 and 0.5. This indicates that reading proficiency is positively correlated to mathematics in general, and that it may play a more important part in some items than others. Twenty items showed a correlation coefficient of 0.4 or higher. These items will be subject of further investigation in the following section. It must be noted that these items also showed a relatively high correlation with the total mathematics score, so a general literacy effect is probably also involved.

The mathematics items in PISA can be categorised using several criteria: content, process, situation, and item format. The 20 items that correlated with reading with a coefficient of 0.4 or higher will be examined and compared with the total number of mathematics items, according to these criteria. They will also be analysed with regard to difficulty (threshold level) and to the quantity of text that

students have to read to solve the mathematical problem. Finally, a selection of highly correlating texts will be examined in terms of sentence complexity, use of low frequency words, inferences needed to understand the meaning of the task, redundant and/or misleading information, illustrations and figures connected to the text, and the necessity of previous knowledge. For comparison we will look at some general characteristics of the items that showed the lowest correlation with reading to see in what way they differ from the highly correlating items.

The following figures show the percentage distribution of highly correlating items compared to all items with regard to *content*, *process* and *item format*.

Figure 1 *Percentage distribution of highly correlating items compared to all items with regard to content*
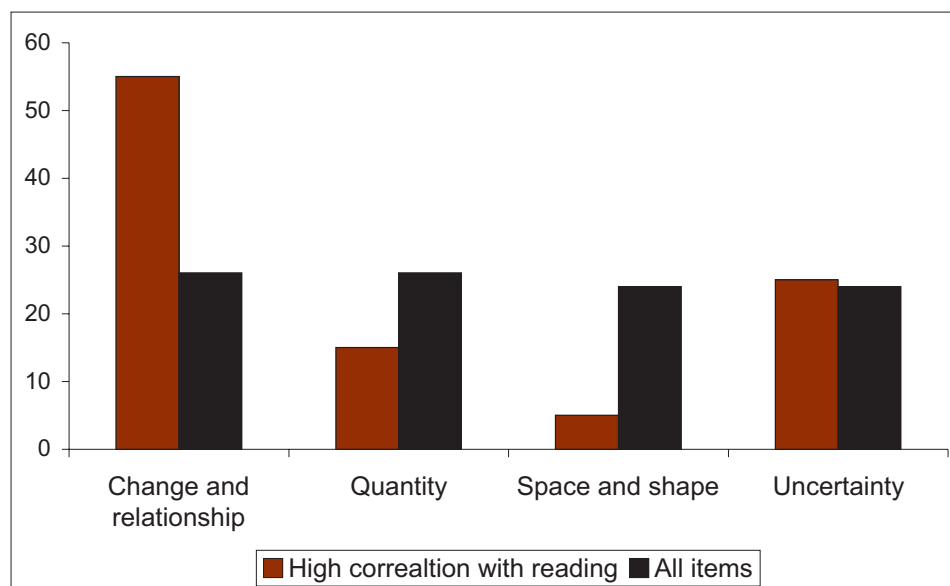
The figure shows that mathematical *content* is evenly distributed in the PISA test as a whole. This is not the case for the highly correlating items. The theme of *change and relationship* is clearly overrepresented, and *space and shape* is correspondingly underrepresented among highly correlating items. According to the PISA mathematical framework (OECD 2004) the theme of *change and relationships* involves mathematical manifestations of change and of functional relationships and dependencies among variables. Relationships are given a variety of different representations, which may serve different purposes and have different properties. Translation between representations is often of key importance in dealing with

situations and tasks. The theme of *space and shape* relates to spatial and geometrical phenomena and relationships (geometry). It involves looking for similarities and differences when analysing the components of shapes, as well as understanding the properties of objects and their relative positions.

Change and relationships tasks definitely require understanding and interpretation of verbal expressions and in some cases the ability also to express this understanding verbally. *Space and shape* tasks, however, require more technical knowledge and understanding, and do not seem to rely so much on reading comprehension.

For *process* tasks there are no significant differences in percentage distribution between all items and highly correlating items. *Item format* tasks*, however, contain a significant overrepresentation of open response items among the highly correlating items and a corresponding underrepresentation of multiple choice items. This is likely to be related to the students' language production, which again may be closely related to reading comprehension. Writing skills and the ability and motivation to write a long and detailed answer may be of importance here.

Figure 2 *Percentage distribution of highly correlating items compared to all items with regard to process*
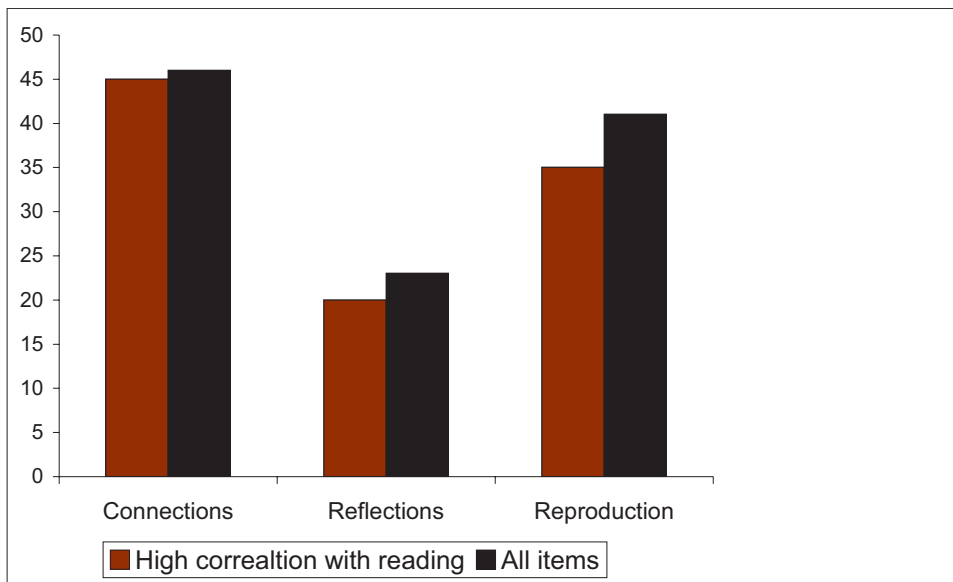
Figure 3 *Percentage distribution of highly correlating items compared to all items with regard to item format*
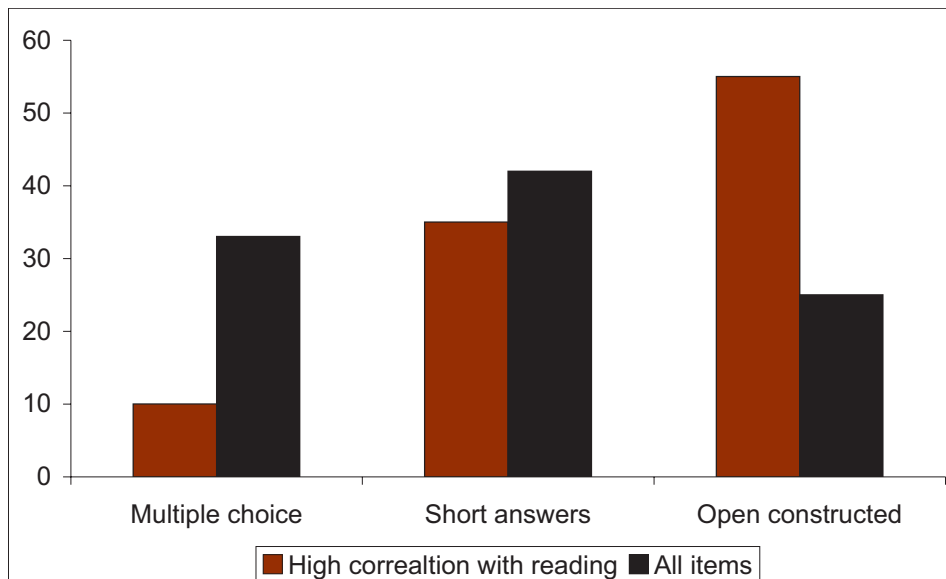
Difficulty

Highly correlating items are on average more difficult than the maths items overall. The average threshold levels for the first score point are 558 for highly correlating items and 545 for all items. The PISA 2003 results for Norwegian and Swedish students showed that on average 46 % of answers were correct for highly correlating items and 50% of answers were correct for all items. A question arises: Are the highly correlating items more difficult because they are more demanding to read? As mentioned earlier there may also be a general literacy factor involved. One possible way of answering this question would be to rewrite some of the more complicated items using simpler words. A substantial increase in the number of correct answers would lend support to the hypothesis that a general literacy factor was involved in many of the difficulties originally connected with these items. No change in the number of correct answers would indicate that a general literacy factor could be ruled out as the cause of difficulties in answering the questions.

Text length

Each item in PISA starts with an introduction. The introduction is either a separate paragraph introducing the item, or an integral part of the item. The length of the introduction in the 84 PISA items ranges from 8 to 278 words with an average of 74 words. One might expect that the items that have high correlation with reading

would contain more introductory text than the average. This was not confirmed by our analysis. The average number of introductory words for the 20 highly correlating items is 77. Thus, reading comprehension in mathematics must have more to do with textual characteristics than with the number of words.

Textual qualities

In the following section we present an analysis of the textual qualities of five of the 20 highly correlating items. It must be emphasised that the relevance of the mathematical knowledge needed to solve the tasks will not be discussed.

Walking[2]

In the introduction to the item on walking the problem is presented in an abstract way using mathematical symbols. At the stage when the formula $\frac{n}{P} = 140$ is presented, the symbol P has already been introduced, whereas the symbol n has not yet been explained. Thus, the reader may become confused on seeing the n, not knowing what it means. For poor readers this confusion may lead to frustration and in some cases to a total comprehension breakdown.

Question 01 The question starts with a hypothetical statement: "*If* the formula…" Thus, the formula is immediately questioned, which might lead the reader to think that he or she is expected to judge the validity of the formula.

Question 03 In contrast to the hypothetical "*If*…." in question 01, this question states that the formula actually applies to Bernard's walking. This may reinforce the confusion about the validity of the formula in question 01.

Growing up

The introduction to the item on growing up contains one sentence only. However, it is loaded with detailed information. Below the introduction is a graph with two axes and two curves, which students have to interpret to understand the task.

Question 03 This task requires a written explanation of "(…) how the graph shows that on average the growth rate for girls slows down (…)". The word "average" might lead to the assumption that there has to be some calculation of a mean value. Although the word "average" is consistent with the information in the graph, and with the text in the introduction, it may affect the reader's comprehension and lead him or her astray.

---

2. The question number corresponds to the number given to the item in the PISA 2003 assessment

Support for president

Question 01 The introduction to this item is part of the question. The total length of text is 136 words (two times the average text length), and it contains a number of low frequency words and expressions, i.e. *opinion polls, forthcoming, newspaper publishers, nationwide, randomly and predicting*. Furthermore, there are two problems connected to the percentages given for each newspaper: Firstly, it is not explicitly stated what they represent. Secondly, the percentage information is redundant in the context of the question. On seeing detailed percentage figures in a mathematical task, poor readers may think that the figures are there to be used in solving the task. Finally, this is an open question where students have to give two reasons to support their answer in written form. This item requires significant knowledge of social science. It has to be questioned what is actually being measured here.

Robberies

The problem presented in this task is based upon a lot of implicit information. First of all students have to interpret the reporter's expression "a huge increase" and the expression "reasonable interpretation" given in the question, and then they need to understand that the relation between the two expressions is the main focus of the question. Secondly they must interpret the graph and understand that the lengths of two graphs cannot be compared directly from their visual appearance. They have to construct an imaginary picture of the full length of the staples to understand that the actual difference between the staples is not significant, and thus draw the conclusion that the reporter's statement is not reasonable. This is also an open response item that requires the students to write an explanation to support their answer.

Finally we looked at the items with the lowest correlation with reading and found the following general characteristics:
1. Straightforward language
2. No difficult or low frequency words
3. Number of words below average
4. Multiple choice or short answer questions.

From the textual analyses of the five highly correlating items above, we assume that the following reading comprehension problems may influence mathematics performance in PISA:

Problems related to the student:
1. Comprehension of mathematical symbols
2. Interpretation of charts and graphs
3. Ability to express and explain the answer in own written words
4. Lack of relevant background knowledge

Problems related to the text:
1. Implicit information that requires inferences and interpretations of abstract relations
2. Misleading information
3. Low frequency words and expressions

Conclusions

To summarise, this study of the results in mathematics and reading from 4595 students who participated in PISA 2003 in Sweden and Norway shows that reading proficiency is positively related to mathematics in general and plays a more important part in some items than in others. Twenty maths items with a correlation to reading of 0.40 or higher were examined and compared with the total number of mathematics items in relation to content, process, situation, item format, difficulty and the quantity of text that students have to read to solve the mathematical problem. Concerning *content* it was found that items on the theme of change and relationships were clearly overrepresented and those on the theme of space and shape were correspondingly underrepresented among highly correlating items. For *process* and *situation* tasks there were no significant differences between all items and highly correlating items in terms of percentage distribution. *Item format* tasks contained a significant over-representation of open response items among the highly correlating items and an underrepresentation of multiple choice items. Highly correlating items were on the average more *difficult* than the math items overall. There was no difference in the number of *introductory words* that students were required to read to solve the mathe-matical problem between highly correlating items and all items. Thus, the influence of reading proficiency on students' ability to solve maths items was obvious in items on the theme of change and relationships, and in open response items and more difficult items. We then analysed five highly correlating items with regard to textual qualities. It was found that a lack of relevant background knowledge, comprehension of mathematical symbols, charts and graphs and the ability to express and explain answers in their own written words were examples of problems that might influence students' mathematics performance. Problems related to the text that could have a negative influence on mathematics performance were implicit information that requires inferences and interpretations of abstract relations, misleading information and low frequency words and expressions.

In short, the results show that there are significant connections between students' mathematical literacy and their reading ability. Furthermore, some mathematical tasks, e.g. items requiring long and detailed written answers and items that are on average more difficult than the maths items overall, show higher correlations with reading than others. The texts contained in the highly correlating items are no longer than the items overall, but they seem to require more understanding and interpretation of verbal expressions than the texts contained in the weakly correlating items. Thus, our study confirms findings by Otterburn and Nicholson (1976) and Shuard and Rothery (1988) about students' problems with mathematical concepts and words and by Åberg-Bengtsson (1994) concerning students` problems with interpreting graphic illustrations. Furthermore, our study supports findings of a relation between ability to solve mathematical problems and text comprehension (i.e. Niss & Højgaard Jensen, 2002; Möllehed, 2001). As suggested by Chinn and Ashcroft (1998), some mathematical texts may also be misunderstood by poor readers or by students who don't read accurately. These findings together indicate that reading literacy plays an important part in the understanding and interpreting of mathematical tasks.

On the basis of these findings we suggest that mathematics teachers should pay more attention to the teaching of reading comprehension in mathematics. It is essential that students understand early that the use of mathematical symbols is a way to express meaning (Mellin-Olsen, 1984). According to Sterner and Lundberg (2002) it is important that students speak, read and write about mathematics. Thoughts and ideas become more visible for reflection and thought when they are transformed into words. This will lead to deeper understanding. Authors of textbooks in mathematics, as well as teachers, who write mathematics tests for their students, should have good language and communication competencies.

In conclusion, we provide some simple pedagogical advice based on available research on the relation between reading proficiency and achievements in mathematics.

Make sure that students

- understand newly introduced mathematical concepts
- understand that some common words have a different meaning when used in a mathematical context
- understand how diagrams, tables and charts in mathematical tasks should be interpreted
- are given many opportunities to reflect on different possible solutions to mathematical problems both orally and in written form
- are given opportunities to practice expressing their answers to mathematical problems in complete sentences both orally and in written form

- who are poor readers have been able to decode and understand all the words in a mathematical task before starting to solve it
- are using textbooks with good language quality in mathematics.

References

Ahlström, R. (2001). Variabler och mönster. *Nämnaren* 28 (1), 27-31.

Chinn, S. & Ashcroft, R. (1998). *Mathematics for dyslexics: a teaching handbook*. London: Whurr Publishers.

Cowen, C.C. (1991). Teaching and testing mathematics reading. *American mathematical monthly*, 98 (1), 50-53.

El-Naggar, O. (1996). *Specific Learning Difficulties in Mathematics: a Classroom Approach*. Tamworth, Staffs: NASEN.

Fuentes, P. (1998). Reading comprehension in mathematics. *Clearing house*, 72 (2), ss.81-88.

Hubbard, R. (1992). Writing Humanistic Mathematics. *Network Journal*, 7, 81-88.

Jakobsson- Åhl, T. (1999). *Communicating mathematically*. Luleå: Institutionen för kommunikation och språk, Luleå tekniska universitet. (D-uppsats/Luleå tekniska universitet; 1999:12).

Kibel, M. (1992). Linking language to action. In T.R. Miles & Miles (eds.), *Dyslexia and mathematics*. London: Routledge, 42-57.

Mellin-Olsen, S. (1984). *Eleven, matematikken og samfundet: en undervisningslære*. Rud: NKI-forlaget.

Miles, T.R. (1992). Some theoretical considerations. In T.R. Miles & E. Miles (eds.), *Dyslexia and mathematics*. London: Routledge.

Möllehed, E. (2001). *Problemlösning i matematik: en studie av påverkansfaktorer i årskurserna 4-9*. Malmö: Institutionen för pedagogik, Lärarhögskolan.

Niss, M. & Højgaard Jensen, T. (red.) (2002). *Kompetencer og matematikinlaerning – ideer og inspiration til udvikling af matematikundervisning i Danmark*. Kobenhavn: Undervisningsminsteriets forlag. Tillgänglig (2004-11-12): http://pub.uvm.dk/2002/kom/hel.pdf

OECD, (2003). *The PISA 2003 Assessment Framework Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. Paris: OECD publication.

OECD (2004). *PISA 2003: Learning for Tomorrow's world. First results from PISA 2003*. Paris: OECD Publication.

Otterburn, M.K. & Nicholson, A.R. (1976). The language of mathematics. *Mathematics in school* 5 (5), 18-20.

Perera, K. (1980). The assessment of linguistic difficulty in reading material. *Educational review 3*2 (2), 151-161.

Shuard, H. & Rothery, A. (eds.) (1988). *Children reading mathematics*. London: Murray.

Sterner, G. & Lundberg, I. (2002). Läs- *och skrivsvårigheter och lärande i matematik. NCM-Rapport 2002:2.* Göteborg: Nationellt Centrum for Matematikutbildning.

Åberg-Bengtsson, L. (2000). Att bygga, rita och tolka cirkeldiagram. *Nämnaren* 27 (3), 24-29.

Österholm, M. (2004). *Läsa matematiska texter: Förståelse och lärande i läsprocessen.* Linköping Studies in Science and Technology. Theses No. 1134. Linköping: Matematiska institutionen, Linköpings universitet.

Chapter 11

What Lies Behind Low Reading Literacy Performance?

A Comparative Analysis of the Finnish and Swedish Students

Pirjo Linnakylä, Antero Malin and Karin Taube

Abstract

This article analyses the main socio-cultural determinants of low reading literacy achievement among Finnish and Swedish students by exploiting PISA 2003 data. In Finland, Swedish-speaking students were oversampled, which enabled comparative analysis of the Finnish- and Swedish-speaking language groups. To find out which background factors increase the risk of low reading literacy performance, two-level logistic regression modelling was used. The results indicate that the risk of being a low achiever in reading literacy is strongly determined in both coutries by male gender, immigrant status, low socio-economic background, lack of educational and cultural resources at home as well as students' low educational aspiration. The constructed general models were relatively similar for both countries, which lays a solid foundation for joint pedagogic developmental efforts.

Nordic abstracts

Artikkeli perustuu PISA 2003 aineiston jatkoanalyyseihin, joilla selvitettiin PISAn lukukokeissa heikosti menestyneiden suomalaisten ja ruotsalaisten nuorten sosiokulttuurisia taustatekijöitä. Suomessa ruotsinkielisten oppilaiden yliotostaminen mahdollisti eri kieliryhmien välisen vertailun. Tutkimusmenetelmänä käytettiin kaksitasoista logistista regressiomallia, jolla selvitettiin sitä, mitkä taustatekijät kasvattivat riskiä heikkoon lukutaitoon. Tulokset osoittivat, että heikko lukutaito oli yhteydessä sukupuoleen, maahanmuuttajataustaan, sosioekonomiseen statukseen, kodin vähäisiin opiskelua tukeviin ja kulttuurisiin resursseihin sekä nuoren omiin koulutustavoitteisiin. Suomalaisten ja ruotsalaisten nuorten heikon suorituksen riskitekijät olivat lähes samat, mikä innostaa tulevaisuudessa yhteisiin pedagogisiin kehittämistoimiin.

Denna artikel analyserar de viktigaste socio-kulturella faktorerna bakom låga läsprestationer bland finska och svenska elever genom att utnyttja PISA 2003 data. I Finland översamplades svensktalande elever vilket möjliggjorde jämförande analyser av de finsk- och svensktalande språkgrupperna. Med syfte att undersöka vilka bakgrundsfaktorer som ökar risken för låga läsprestationer användes en logistisk regressionsmodell med två nivåer. Resultaten i bägge länderna indikerar att risken att bli lågpresterande i läsning i hög grad avgörs av faktorerna manligt kön, invandrarstatus, låg socioekonomisk bakgrund, brist på utbildningsresurser och kulturella resurser i hemmet liksom elevers låga utbildningsambitioner. De konstruerade generella modellerna var relativt lika för båda länderna, vilket skapar en stabil grund för gemensamma ansträngningar för pedagogisk utveckling.

Introduction

Providing all students with equal access to education and removing obstacles to learning, especially among students from disadvantaged backgrounds, have been leading principles in Nordic education policies. Thus the Nordic school systems are further characterised by the inclusion of special education programmes, the integration of ethnic minorities, and various instructional efforts aimed at minimising low achievement. Historically, the first challenge facing the Nordic countries in their pursuit of equal opportunities in education was geographic barriers. Next, equality was called for different socio-economic groups, then for gender and finally for immigrant students (Husén, 1974; OECD, 2001; Fredriksson, 2002; Lie et al., 2003; Linnakylä et al. 2004).

In recent decades the aspiration towards social justice through education has been accompanied by neo-liberalistic views, which call for ever higher standards in education and for efficiency and effectiveness in school management (Rinne et al. 2000). Schools have started to compete in terms of special profiles, pedagogical innovations and learning achievements. The trend towards increased autonomy granted to schools, along with special curricular programmes, the establishment of private schools, and increased external funding, has also raised concerns about deterioration in Nordic educational equity (Söderberg, 2001). Parents can now select their children's schools because of new ideas based on parents as 'customers' (Brown, 1990). In choosing a school both the quality of education and the social climate of the school have been stressed. This has often meant abandoning the system of school districts and a uniform curriculum typical of the policy of equity. Those arguing for greater parental choice emphasise consideration of individuality and talent, but forget that not all parents have the same opportunities, education or wealth, to make such choices. Parents with the highest levels of education and wealth are best acquainted with the alternatives and most interested in supporting

their children's education and choosing a school for them (Söderberg, 2001; Lie et al. 2003). In addition, a family's social network may serve as a channel for information about the best schools, course options, and further educational opportunities. Less well-educated parents tend to be less concerned about school issues. The parents' educational background, occupation, and related economic status also have a bearing on a family's educational and cultural resources beyond school. Hence, the economic, cultural and social capital of the family influences the children's learning in various ways, either promoting or hindering it (Bourdieu, 1986).

How successful have the Finnish and Swedish school systems been at correcting imbalances in literacy learning due to individual, linguistic and ethnic, socio-economic and cultural differences? Which socio-cultural factors increase the risk of low performance? This article presents secondary analyses of the Finnish and Swedish PISA 2003 data aimed at identifying some significant factors associated with low performance in reading literacy.

Aims of the study

The aim of the present study is to examine, contrast and compare some individual, socio-economic and cultural determinants of low reading literacy performance among Finnish and Swedish students by exploiting the PISA data collected in 2003 from representative samples of 15-year-old Finnish and Swedish students (OECD, 2004). In Finland, the Swedish-speaking students were oversampled, which enabled comparison of the two language groups, Finnish and Swedish, to be made. Since mathematics was a major domain in PISA 2003, not all students answered the reading literacy tasks, and thus the final sub-sample used for the analysis is smaller than the entire PISA 2003 sample. In Finland 3144 students answered the reading literacy tasks and in Sweden the number was 2443. Among the Finnish students, 2490 attended Finnish-speaking and 654 Swedish-speaking schools.

The main questions addressed in this study were as follows:

- Which background factors are the strongest determinants of low reading literacy performance, as compared and contrasted with the background factors related to a sufficiently high performance in Finland and Sweden?
- To what extent do the risk factors differ in Finland and Sweden; and what are the similarities?
- What similarities and differences can be found in the factors responsible for low achievement by Finnish- and Swedish-speaking students in Finland?
- What such similarities and differences can be found for Swedish students and Swedish-speaking Finnish student?

A limitation of this study is that the analyses undertaken are restricted to the variables available. Thus, although the PISA 2003 data are relatively rich in terms of student characteristics and home background factors, the learning environment and instructional strategies of reading literacy were not addressed in detail in PISA 2003, when mathematics was the main assessment domain of the study. Nevertheless, students' interest and engagement in reading activities in their free-time was assessed in Finland as a national option, which enables more varied analyses of the Finnish data.

Data and method

In the initial analyses of the PISA 2003 data, students were assigned to five levels of performance according to their reading literacy scores. Students at the three highest performance levels (5, 4 and 3) demonstrated reading literacy skills considered sufficiently high for further studies, working life and active citizenship. Students proficient at Level 2, 1 and below were, instead, considered to need additional or special reading literacy education in order to cope with further studies and working life.

The initial results of PISA 2003 showed that Finnish and Swedish students are among the best readers across all OECD countries (OECD, 2004). Even though the national mean performances of Finland and Sweden proved high, there were also low-performing students in both countries (Figure 1).

In the OECD countries, on average, 58% of students achieved a sufficiently high proficiency level (Levels 5, 4 and 3). In Finland this level was reached by 77% of students and in Sweden by 67% of students. Thus, the OECD countries had an average of 42% of students performing at Levels 2, 1 and below, while the corresponding figures for Finland and Sweden were 23% and 33% respectively. In Finland, there were relatively more low-performers among Swedish-speaking students (27%) than among Finnish-speaking students (22%).

In this study, risk factors for low reading literacy performance were studied by contrasting the group of low-performing students (at or below Level 2) with that of high-performers (at Levels 5, 4 and 3) by exploiting two-level logistic regression models (Snijders & Bosker, 2002). In addition, the Finnish data were analysed separately for both language groups. The variables included in the analyses are presented in Table 1. The background variables were chosen on the basis of previous findings on significant factors associated with low reading literacy achievement (e.g. Elley, 1994; Linnakylä et al. 2004; Thorndike, 1973).

Figure 1 *Percentage of students at low- and high performance levels in reading literacy*
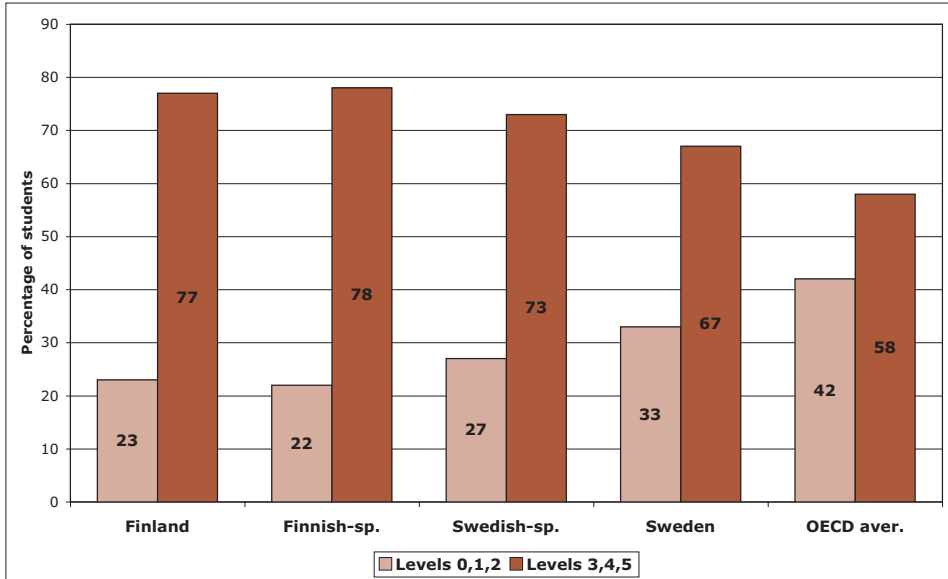
The outcome variable describing group membership was dichotomous: the student belonged either to the group of low-performers (1) or to the reference group of high-performers (0). The performance variable was based on weighted likelihood estimates (WLE) in the PISA combined reading literacy scale score (for details, see OECD, 2001). The performance variables in the models are logit transformations

Table 1 *Variables included in the models*

Outcome variable: reading literacy level based on the PISA 2003 performance scale score	
Explanatory variables:	
Individual:	gender, attitudes towards school, teacher-student relations, educational aspiration, engagement in reading (in Finland)
Socio-demographic:	migration, school language, family structure/ single parent
Socio-economic:	parents' occupational status
Home-cultural:	cultural possessions at home, home educational resources

of students' probability of belonging to the group of low-performers. The coefficient b values, associated with p-values indicate the statistical significance of the estimates and show the effect of the background factors on the logit. The constant is the logit of those students all of whose background factors are equal to zero. The estimated odds ratios Exp(b) are multiplying factors by which a student's risk of belonging to the group of low-performers increases or decreases if the student has a given characteristic.

Results

The results of the two-level models are first presented separately for Finland and for Sweden in order to investigate to what extent the same or different factors are involved in increasing or decreasing the risk of low performance. The Finnish data comprise both Finnish- and Swedish-speaking students. The data are weighted in relation to the ratio of Swedish-speakers (5.8% of the population). Secondly, the Finnish model is restructured by including the variable *engagement in reading* in the model. Thirdly, the data of the Finnish- and Swedish-speaking students in Finland are fitted into the wider models with the *engagement in reading* variable. The data are then weighted in relation to the actual sample sizes of the language groups. Finally, the respective models for the Swedish-speaking students in Finland and for the Swedes are presented and compared.

Risk factors for low reading literacy performance in Finland and Sweden

Those explanatory variables that were statistically significant in decreasing or increasing the risk of low reading literacy performance are shown in boldface in Table 2.

The coefficients obtained through the logistic regression models indicate that the factors increasing students' risk of low performance in reading literacy are largely similar in Finland and Sweden. In both countries *gender* and *immigrant status* were strongly associated with low performance. Among boys the risk of low performance was doubled in both countries. The risk of low performance was even more pronounced for immigrant students. In Sweden the risk was also higher among non-native students than among students who were born in Sweden but whose parents were foreign-born. In Finland there were so few first-generation immigrant students in the sample that they were embedded in the native population. The risk of low performance among non-natives in Finland was about nine times higher than among other students. In Sweden, the risk was also high for the immigrant groups; among non-natives it was tripled and among first-generation immigrant students it was doubled.

Table 2 *Risk factors of low reading literacy performance in Finland and Sweden*

	Finland			Sweden		
Fixed effects	*b*	*p*	Exp(*b*)	*b*	*p*	Exp(*b*)
Constant	**-0.69**	0.000	**0.50**	0.11	0.512	1.12
School language (Swedish)	0.23	0.245	1.26	-	-	-
Gender (male)	**0.85**	0.000	**2.35**	**0.55**	0.000	**1.74**
First generation students	-	-	-	**0.48**	0.023	**1.61**
Non-native students	**2.18**	0.000	**8.87**	**1.09**	0.000	**2.96**
Single-parent family	-0.12	0.289	0.88	0.06	0.578	1.07
Socio-economic status	**-0.29**	0.000	**0.75**	**-0.18**	0.001	**0.84**
Cultural possessions at home	**-0.19**	0.000	**0.82**	**-0.21**	0.000	**0.81**
Attitudes towards school	**-0.17**	0.003	**0.85**	-0.11	0.051	0.89
Home educational resources	**-0.13**	0.013	**0.88**	**-0.17**	0.001	**0.85**
Teacher-student relations	-0.04	0.458	0.96	**-0.23**	0.000	**0.80**
Educational aspirations	**-0.29**	0.000	**0.75**	**-0.36**	0.000	**0.70**
Random effects						
Between-school	0.15			0.12		
Within-school	0.94			0.95		
N of schools	197			185		
N of students	3036			2307		
% of correctly classified students	79.1			73.2		

Note: Statistically significant coefficients at p ≤ 0.05 are indicated in bold typeface

Additionally, students from a *lower socio-economic background*, measured as parents' occupational status, faced a higher risk of low achievement than others. The risk was also increased by the family's *lack of possessions related to classical culture*, such as literature, books of poetry and works of art. The risk was further increased if the family was *lacking home educational resources*, such as a quiet place and a desk for studying, a dictionary, a calculator and books to help with school work.

Apart from the family-related factors, students' personal characteristics, such as *educational aspiration* and *attitudes toward school and learning*, were associated with their reading literacy performance. In both countries the risk was higher for students who did not expect to spend much time on further studies after finishing their compulsory basic education. Furthermore, the risk was increased in Finland if student's attitudes towards school were negative; i.e. if the student felt that school had done little to prepare him/her for adult life; had been a waste of time; had not helped him/her to gain confidence to make decisions; and had not taught things which could be useful in a job. In Sweden, on the other hand, negative attitudes

towards school did not increase the risk. But the risk was increased in Sweden if students' relationship with teachers was problematic, whereas this was not a risk factor in Finland.

A comparison of the two-level models indicates that the school-level residual variance was slightly larger in Finland than in Sweden and that Finland thus showed larger *between-school variance*.

Engagement in reading in the restructured Finnish model

The item *engagement in reading* was measured in Finland as a national option in PISA 2003, because in previous analyses based on the PISA 2000 data this factor had been found to be most significant in explaining both reading literacy performance and gender difference (Linnakylä & Malin, 2003; Linnakylä et al. 2004). In this study, again, the coefficients obtained using the expanded Finnish

Table 3 *Risk factors of low reading literacy performance in Finland, including engagement in reading*

	Finland		
Fixed effects	*b*	*p*	Exp(*b*)
Constant	**-0.62**	0.001	**0.54**
School language (Swedish)	0.25	0.215	1.28
Gender (male)	**0.32**	0.003	**1.38**
Non-native students	**2.11**	0.000	**8.27**
Single-parent family	-0.12	0.317	0.89
Socio-economic status	**-0.29**	0.000	**0.75**
Cultural possessions at home	**-0.12**	0.027	**0.89**
Attitudes towards school	**-0.13**	0.033	**0.88**
Home educational resources	-0.09	0.087	0.91
Teacher-student relations	0.04	0.487	1.04
Educational aspirations	**-0.26**	0.000	**0.77**
Engagement in reading	**-0.62**	0.000	**0.54**
Random effects			
Between-school	0.16		
Within-school	0.92		
N of schools	197		
N of students	3007		
% of correctly classified students	79.7		

Note: Statistically significant coefficients at p ≤ 0.05 are indicated in bold typeface

model including the engagement variable indicate that positive attitudes and active engagement in reading decreased the risk of low performance significantly (Table 3). Furthermore, the impact of gender decreased significantly in the expanded model, which indicates that evening out engagement in reading activities could considerably reduce the gender difference in reading performance. In addition, the impact of home educational resources lost its significance in the reconstructed model.

These findings indicate that students, mainly boys, who were not interested and actively engaged in reading had a significantly higher risk of low performance. These students reported that they read only when they had to; that reading was a waste of time; that they read only to get information; that reading was not one of their favourite hobbies; that they did not like talking about books; that it was hard for them to finish a book; that they did not feel happy if they received a book as a present; that they did not enjoy going to a bookstore or a library; and that they could not sit still and read for more than a few minutes.

Differences between Finnish- and Swedish-speaking students in Finland

The models were next fitted separately to the Finnish- and Swedish-speaking students' data to find out to what extent similar factors increased the risk of low performance in the two language groups in Finland.

The coefficients of the two separate models of Finnish- and Swedish-speaking Finns (Table 4) – including *engagement in reading* data – indicate that the factors increasing students' risk of low performance in reading literacy were surprisingly different in the two language groups in Finland. Only students' active engagement in reading seemed to similarly decrease the risk of low performance in both groups. Otherwise the significant risk factors differed. In the Finnish-speaking students' model, gender, immigrant status, socio-economic background, cultural possessions and educational aspiration were critical factors. Boys and particularly non-native students were at a greater risk than girls and native students. Furthermore, students from a *lower socio-economic background*, particularly, if the family was lacking of possessions related to classical culture, were at a greater risk of low performance than others. In addition, low educational aspiration increased the risk of low performance. The Swedish-speaking students' risk of low performance was, in contrast, significantly associated with negative attitudes towards school and lack of educational resources at home. In the Swedish-speakers' expanded model, the risk of low performance among the boys was no longer significant. This is mainly due to the fact that the performance in reading of Swedish-speaking girls' was significantly lower than that of Finnish-speaking girls' and thus the gender

Table 4 *Risk factors of low reading literacy performance of Finnish- and Swedish-speaking Finns*

Fixed effects	Finnish-speaking Finns			Swedish-speaking Finns		
	b	*p*	*Exp(b)*	*b*	*p*	*Exp(b)*
Constant	**-0.58**	0.004	**0.56**	**-0.93**	0.009	**0.40**
Gender (male)	**0.34**	0.005	**1.41**	0.13	0.535	1.14
Non-native students	**2.17**	0.000	**8.77**	1.02	0.251	2.76
Single-parent family	-0.14	0.293	0.87	0.35	0.211	1.42
Socio-economic status	**-0.31**	0.000	**0.74**	-0.16	0.117	0.85
Cultural possessions at home	**-0.12**	0.046	**0.89**	-0.14	0.259	0.87
Attitudes towards school	-0.11	0.099	0.90	**-0.39**	0.002	**0.68**
Home educational resources	-0.07	0.302	0.94	**-0.37**	0.000	**0.69**
Teacher-student relations	0.04	0.591	1.04	0.10	0.419	1.11
Educational aspirations	**-0.27**	0.000	**0.76**	-0.11	0.184	0.89
Engagement in reading	**-0.62**	0.000	**0.54**	**-0.63**	0.000	**0.53**
Random effects						
Between-school	0.12			0.18		
Within-school	0.94			0.96		
N of schools	147			50		
N of students	2368			639		
% of correctly classified students	80.6			78.2		

Note: Statistically significant coefficients at p ≤ 0.05 are indicated in bold typeface

difference was smaller among the Swedish-speakers, particularly in the group that was not actively engaged in reading activities. But if *engagement in reading* was not controlled for, the risk of low performance appeared significant among the Swedish-speaking boys as well (Table 5).

Comparing Swedes and Swedish-speaking Finns

The models fitted separately to the data for Swedish students and Swedish-speaking Finnish students (Table 5) – without *engagement in reading* – gives an opportunity to compare the similarities and differences in risk factors among Swedish-speakers in the two countries with two different school systems. Similarly and significantly in both countries, the risk of low performance was increased by male gender, as well as by lack of cultural possessions and educational resources at home. Otherwise, the models differed. In Swedish-speaking schools in Finland, immigrant status was not significantly associated with low performance, which may be largely related to the fact that there are very few immigrant students in the Swedish-

Table 5 *Comparing Swedes' and Swedish-speaking Finns' risk factors*

Fixed effects	Sweden			Swedish-speaking Finns		
	b	*p*	*Exp(b)*	*b*	*p*	*Exp(b)*
Constant	0.11	0.512	1.12	**-0.89**	**0.014**	**0.41**
Gender (male)	**0.55**	**0.000**	**1.74**	**0.55**	**0.005**	**1.74**
First generation students	**0.48**	**0.023**	**1.61**	-	-	-
Non-native students	**1.09**	**0.000**	**2.96**	1.16	0.185	3.18
Single-parent family	0.06	0.578	1.07	0.36	0.197	1.43
Socio-economic status	**-0.18**	**0.001**	**0.84**	-0.15	0.152	0.86
Cultural possessions at home	**-0.21**	**0.000**	**0.81**	**-0.24**	**0.050**	**0.79**
Attitudes towards school	-0.11	0.051	0.89	**-0.44**	**0.001**	**0.65**
Home educational resources	**-0.17**	**0.001**	**0.85**	**-0.40**	**0.000**	**0.67**
Teacher-student relations	**-0.23**	**0.000**	**0.80**	0.01	0.923	1.01
Educational aspirations	**-0.36**	**0.000**	**0.70**	-0.15	0.070	0.86
Random effects						
Between-school	0.12			0.20		
Within-school	0.95			0.99		
N of schools	185			50		
N of students	2307			640		
% of correctly classified students	73.2			77.2		

Note: Statistically significant coefficients at p ≤ 0.05 are indicated in bold typeface

speaking schools. If there are immigrants, they are usually Swedish speakers returning from Sweden. Furthermore, low socio-economic background was associated with low performance only in Sweden. The two models also differed in attitudinal factors. Positive attitudes towards school decreased the risk among Swedish-speakers in Finland, but were insignificant in Sweden, where, again, students' positive relations with teachers decreased the risk. The risk was further decreased by students' high educational aspirations in Sweden, which was not the case among Swedish-speaking Finns.

Conclusion and discussion

The findings reveal that the factors increasing students' risk of low reading literacy performance are, in many respects, similar in Finland and in Sweden. Immigrant status, male gender, low socio-economic background, lack of educational and cultural resources, and low educational aspiration were all factors that increased the risk of low performance in both countries, when the other factors were simultaneously controlled for. Irrespective of male gender, immigrant status was a

very strong risk factor. In Sweden, the distinction was made between non-native and first generation immigrants. It was evident that the risk of low performance was weaker among first generation Swedes than among non-native students. In Finland, the number of first generation immigrant students was so small that they were included with the native Finns. The risk of low performance among non-natives in Finland was very high, about nine times higher than among other students. In Sweden the risk was also high for the immigrant groups, although not as striking; among non-natives it was tripled and among first-generation immigrant students it was doubled.

Attitudinal factors had somewhat different tendencies to increase the risk. In Finland, the risk was increased if the students' attitudes towards school were negative in terms of their future studies and adult life. This was particularly the case in the Swedish-speaking group in Finland. In Sweden, on the other hand, the risk was increased if students' relationships with teachers was problematic. This was not the case in Finland. Including an attitudinal variable *engagement in reading* into the model changed the structure of risk factors to some extent. The impact of gender decreased dramatically and indicated that by equalising engagement in reading in both gender groups, the performance gap could be narrowed. Among the Swedish-speaking Finns the risk associated with male gender disappeared when the engagement factor was included. Likewise, the impact of cultural possessions at home lost its significance. Among the Finnish-speakers, the impact of the gender remained significant but was considerably reduced.

Comparing and contrasting these models reveals that the respective models for the Swedes and the Swedish-speaking Finns, as well as the models of the two language groups in Finland, differed more from each other than the first general models for the two countries. In each model, differences between schools had little effect on the risk of low reading literacy performance. The risk factors were mainly those within school and between students. Although in Finland there was slightly more variation between schools in terms of students falling into the risk group.

In the PISA studies, the Finnish and Swedish school systems have been found to be quite successful in producing learning outcomes of both high quality and high equity. The results of this further study, however, indicate that there is still a lot to do to correct imbalances in literacy learning due to individual, ethnic, language, socio-economic and cultural differences. In both countries, students with immigrant backgrounds, particularly males with low educational aspirations, who come from families with low socio-economic status and from homes lacking educational and cultural resources, are disadvantaged. These findings suggest that poor readers need affective, cognitive and social support. Such triple-based support

can be enhanced through the students' curiosity or desire to learn about certain topics, enjoyable real-world interaction, interesting and exciting texts, personal choice of reading materials, significant literary experiences and collaboration with peers (Guthrie et al. 1996; Guthrie & Wigfield, 2000).

Within a supportive school environment the importance of listening to the voices of students must not be underestimated (McCombs, 1996). When students are asked what is right about schools, they most frequently mention social relationships in which people care, listen, are honest and open, understand, and respect others, including low achievers (Poplin & Weeres, 1993). Further analyses of the PISA 2000 data also revealed that students' relationships with their teachers had a strong correlation with their engagement in reading, particularly in Nordic countries (OECD, 2002). In Sweden, a lack of confiding relationships between students and adults at school has been found to be related to students' difficulties in completing compulsory education with satisfactory marks, which strongly predicts young people's opportunities in the labour market and in society later in life (National Agency for Education, 2001).

Although students' socio-economic, ethnic and cultural home background as such cannot be changed in the short term, low achievers' parents could become more involved in their children's school. Home-school partnerships can have a positive effect on literacy, if families and teachers together develop ways of communicating and building meaningful curricula that extend the insular classroom community. The key elements of reciprocity and respect must be jointly constructed by parents and teachers (Baker et al., 1996).

In the light of these findings, fostering boys' interest and engagement in reading seems to pose a special challenge for education, since boys were clearly overrepresented among low achievers. If the attitudes of low-achieving boys towards reading are indeed so negative that they do not read anything unless they have to and that they consider reading just a waste of time, there is certainly need for a cultural change. In pedagogical terms this means that we should invest heavily in motivational development. Teachers should have a knowledge of the kinds of literature and reading materials that boys also find interesting, such as science fiction and fantasy stories (Baker et al. 1996).

A recent national study in Sweden has shown that schoolwork has increasingly become the individual student's own responsibility (National Agency for Education, 2004). This tendency might be a further disadvantage for boys, because of working hard at school is not seen as compatible with being masculine (Pickering, 1997). Thus, the individualisation of schoolwork might have a negative

influence on boys with little or no motivation for reading and studying. We must realise the importance of confronting existing anti-study school cultures, if we want to prevent their negative impacts on learning (Björnsson, 2005).

The pursuit of equal opportunities for learning has been put to a severe test in Sweden and, more recently, in Finland, due to increasing numbers of immigrant students and growing cultural heterogeneity. This presents a special challenge for reading literacy education, and the findings of this study attest that there still remains much to do if we want to ensure that immigrant students have equal opportunities to learn literacy, ideally both in their own mother tongues and in the languages of their current home countries. Immigrant background is by no means homogeneous and includes a variety of native countries, mother tongues and cultures. The immigrant sample in PISA 2003 data for Sweden and particularly for Finland was too small to allow comparisons between groups of students with different native countries and mother tongues. Other studies, however, have shown that in Sweden, for example, the reading performance of German- and English-speaking immigrant students' is higher than that of Arabic, Turkish, Romany and Somali students, which may be due to various linguistic, cultural, economic, and societal factors (Fredriksson, 2002; Taube & Fredriksson, 1995).

In facing this challenge Finland has a lot to learn from Sweden, where, unlike Finland, there has been ample experience of educating immigrant students' (Fredriksson, 2002). Sweden, however, cannot ignore the challenge either, since immigrant background was found to increase the risk of low achievement, even though socio-economic and cultural background, gender and attitudinal factors were simultaneously controlled for. This suggests that in order to reduce the risk of low-achievement by immigrant students we must look at the problem in a wider context, which would include pedagogical, linguistic, cultural and social support. Each immigrant student has unique background reasons for immigrating. To be able to help and support these students in their new country it is important to know the special needs of each student and his/her family. Only in partnership with the family can the school succeed in equalising opportunities for immigrant students (Fredriksson & Taube, 2003; Linnakylä et al. 2004). Most of the families have moved to a new country to achieve a happier life and equal opportunities in education, work and life as active citizens.

References

Baker, L., Afflerbach, P., & Reingking, D. (1996). Developing engaged readers in school and home communities: An overview, in l. Baker, P. Afflerbach & D. Reinking (Eds.) *Developing engaged readers in school and home communities* (pp. xiii-xxvii). Mahwah, NJ: Erlbaum.

Björnsson, M. (2005). *Gender and success at school: interpretations and perspectives.* Stockholm: National agency for education. Government commission.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (ed.) *The handbook of theory and research for the sociology of education* (pp. 241–259). New York: Greenwood Press.

Brown, P. (1990). The third wave. Education and ideology of parentocracy, *British Journal of Sociology of Education,* 11, 65–84.

Bryk, A. & Raudenbush, S. (1992). *Hierarchical linear models: application and data analysis methods* Newbury Park: Sage.

Elley, W. B. (1994). *The IEA study of reading literacy: achievement and instruction in thirty-two school systems* Oxford: Pergamon Press.

Fredriksson, U. (2002). *Reading skills among students of immigrant origin in Stockholm.* Stockholm: Stockhol University, Institute of International Education.

Fredriksson, U. & Taube, K. (2003). Svenska som andraspråk och kulturmöten (Swedish as the second language and cultural encounters), in: L. Bjar & C. Liberg (Eds.). *Barn utvecklar sitt språk* (pp. 154–172). Lund: Studentlitteratur.

Guthrie, J.T., McGough, K., Bennett, l. & Rice, M.E. (1996). Concept-oriented reading instruction: An integrated curriculum to develop motivations and strategies for reading, in: l. Baker, B. Afflerbach & D. Reinking (Eds.) *Developing engaged readers in school and home communities* (pp. 165–190). Mahwah, NJ: Erlbaum.

Guthrie, J.T. & Wigfield, A. (2000). Engagement and motivation in reading. In M.l. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.) *Handbook of reading research,* Vol. III, (pp. 403–422). New York: Longman.

Husén, T. (1974). *Learning society.* London: Methuen.

Lie, S., Linnakylä, P. & Roe, A. (2003). *Northern lights on PISA. Unity and diversity in the Nordic countries in PISA 2000* Oslo: University of Oslo, Department of teacher education and school development.

Linnakylä, P. & Malin, A. (2003). How to reduce the gender gap in reading literacy. In S. Lie, P. Linnakylä, & A. Roe (Eds.). *Northern lights on PISA. Unity and diversity in the Nordic countries in PISA 2000* (pp. 21–38). Oslo: University of Oslo, Department of teacher education and school development.

Linnakylä, P., Malin, A. & Taube, K. (2004). Factors behind low reading literacy achievement. *Scandinavian Journal of Educational Research 48* (3), 231–250.

McCombs (1996). Alternative perspectives for motivation. in l. Baker, P. Afflerbach & D. Reinking (Eds.) *Developing engaged readers in school and home communities*, (pp. 67–87). Mahwah, NJ: Erlbaum.

National agency for education (2001). *Without complete marks – why do not all students reach the goals?* Stockholm: National agency fo education: Government commission.

National agency for education (2004). *National assessment 2003*. Stockholm: National agency for education.

OECD (2001). *Knowledge and skills for life. First results from pisa 2000*. Paris: OECD.

OECD (2002). *Reading for change. Performance and engagement across countries*. Paris: OECD.

OECD (2004). Learning for tomorrow's world. First results from PISA 2003. Paris: OECD.

Pickering, J. (1997). *Raising boys' achievement*. Stafford: Network Educational Press.

Poplin, M. & Weeres, J. (1993). Listening at the learner's level, *The Executive Educator*, 15, 14–19.

Snijders, T.A.G. & Bosker, R.J. (2002). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London: Sage.

Swedish government official reports (2003). *Young outsiders*. Stockholm: Swedish government official reports, nr. 92.

Söderberg, s. (2001). Attitudes and expectations in relation to school: Swedish findings and some international comparisons. In *Schooling for tomorrow: what schools for the future* (pp. 177–202). Paris: OECD.

Taube, K. & Fredriksson, U. (1995). *Hur läser invandrarelever i Sverige?* Stockholm: Skolverkets rapport nr. 79.

Thorndike, R.L. (1973). *Reading Education in Sixteen Countries: an empirical study*. Stockholm: Almqvist & Wiksell.

Chapter 12

Could Confidence in ICT Boost Boys' Reading Performance?

Kaisa Leino and Antero Malin

Abstract

This article focuses on students' self-reported computing ability - their confidence in tackling Information and Communication Technology (ICT) tasks - and how this relates to reading performance, based on results for Finnish students in PISA 2003. Using a multilevel modelling technique, the effects of several background variables such as gender, socio-economic background and mother tongue (Finnish or Swedish) were controlled for. The association between confidence in ICT and reading literacy performance was clearly stronger for boys than girls. The results suggest that boys seem to benefit from a high level of confidence in routine and Internet tasks and moderate confidence levels in high-level ICT tasks. The girls, on the other hand, benefit from confidence in routine tasks and moderate confidence in high-level tasks. Even though use of ICT and confidence in related ICT tasks explains only a few per cent of variance in reading literacy, it is something we could use to support and encourage particularly boys' literacy practices.

Nordic abstract

Artikkeli tarkastelee oppilaiden käsityksiä heidän omista tietokoneenkäyttötaidoistaan ja tämän tieto- ja viestintäteknisen luottamuksen yhteyttä PISA 2003:ssa mitattuun lukutaidon tasoon. Monitasomalleja hyödyntäen useiden taustamuuttujien, kuten sukupuoli, sosio-ekonominen tausta ja koulun kieli (suomi vai ruotsi), vaikutukset kontrolloitiin. Tulosten mukaan tietoteknisen luottamuksen ja lukutaidon yhteys on selvästi voimakkaampi pojilla kuin tytöillä. Pojat näyttävät hyötyvän, jos heillä on erittäin hyvä itseluottamus tietokoneen rutiini- ja Internet-toimintojen hallitsemisesta sekä ainakin kohtuullinen luottamus myös teknisesti vaativampien toimintojen suhteen. Tytöille taas oli eduksi rutiinitoimintojen hyvä hallinta ja kohtuullinen luottamus myös vaativampien toimintojen hallitsemisesta. Vaikka tietotekniikan käyttö ja luottamus omiin tietoteknisiin taitoihin selittääkin vain muutaman prosentin lukutaidon tason

vaihtelusta, ovat ne kuitenkin tekijöitä, jolla voitaisiin tukea ja rohkaista erityisesti poikien lukuharrastusta ja lukemisen käytänteitä.

Introduction

Today literacy is far more than just reading and writing. Creating, accessing, navigating, evaluating, and problem-solving with different media demand multiliteracy skills and knowledge, practices, important in information society. Computer use, valued very positively by adolescents (e.g. Leino, in press), has a firm place in the everyday literacy practices of most adolescents in Finland, and in other Nordic countries, too. Electronic communication, searching for information, downloading music and programs and playing games were the most popular activities among 15-year olds (e.g. OECD, 2006.)

Various studies have reported that perceived ability to use computers correlates with reading scores (e.g. Leino, 2003; OECD, 2006), but the effect of computers on student achievement depends on the specific ways in which the computers are used (e.g. Fuchs & Woessmann, 2004; OECD, 2006). In particular, use of computers to teach higher-order thinking skills has been shown to be positively related to academic achievement (Wenglinsky, 1998). In results based on data from the PISA 2000 survey, performance in reading increased with the frequency of the use of the Internet and electronic communication (Fuchs & Woessmann, 2004; Leino, 2002), but programming, and using educational software and spreadsheets seemed to have negative relationship with reading (Leino, 2002). In addition, a study by Leino, Linnakylä and Malin (2004) showed that when different kinds of Internet tasks were considered, those students reading a diversity of media texts, 'active multiliterate readers', were nearly as proficient, based on PISA reading scores, as 'active traditional readers' who mainly focused on fiction, newspapers and magazines. Indeed, the literacy practices of 'active multiliterate readers' may prove to be more useful in today's information society. These studies also show that the level of computer use seems not to matter very much, but extremely high levels of use may be counterproductive (e.g. Leino, 2003; OECD, 2006).

Self-efficacy and self-concept beliefs are related to academic achievements: A high degree of confidence anticipates better results. Bandura (1986a) has even claimed that people's behaviour can be better predicted by their beliefs about their capabilities than by their actual accomplishments, for competent functioning requires harmony between self-beliefs on the one hand and possessed skills and knowledge on the other. Self-perceptions of capability help determine what individuals do with the knowledge and skills they have (Bandura, 1986a; Pajares, 2002). This can be seen in our study as a strong correlation between beliefs and actions used.

Self-confidence is related to reading literacy performance and confidence in ICT usage, in particular. The OECD report (2006) showed that confidence in ICT

and performance in mathematics tend to go together. This article takes this observation a step further and disentangles the relationship between computer use and reading literacy by studying whether gender differences in students' confidence in ICT tasks affect reading literacy achievements. Gender differences in reading are especially pronounced in Finland (OECD, 2001). The focus here is on gender differences in students' self-reported confidence in ICT, and how this correlates with reading literacy scores, based on results for Finnish students in PISA 2003. Confidence in ICT tasks is divided into three categories, as specified in the PISA framework: confidence in routine tasks, confidence in Internet tasks, and confidence in high-level tasks (e.g. OECD, 2005, p. 306). In addition, a comparison among Nordic countries for these three categories is presented. The aim of the article is to answer the following questions: Does the students' confidence in ICT tasks relate to the PISA reading literacy scores? Are different categories of confidence in ICT associated with students' reading scores, and are there gender differences? When background variables are controlled for, how does familiarity with computers influence the reading performance of boys and girls?

Self-Confidence in ICT

Confidence is here understood as self-efficacy which as a concept builds on social learning theory. According to Bandura (1986b), "self-efficacy is the belief in one's capabilities to organize and execute the sources of action required to manage prospective situations – an individual's judgement of his or her capabilities to perform given actions". These beliefs influence the choices people make, how they approach new tasks and what kind of outcomes they expect. Self-efficacy revolves around questions of 'can' (e.g. Can I open a file? Can I construct a web page?). (Bandura, 1986b; Pajares & Schunk, 2001.) Even though self-confidence as a learner is an essential factor in motivation and is related to achievements, the researchers are troubled by the chicken-and-egg question of causality (Pajares & Schunk, 2001): for example, where computer use is concerned, confident ICT users usually use computers often and frequent use amplifies skills and confidence.

Real-world experiences and 'hands-on' tasks are extremely motivating. An integrated curriculum including reading and 'hands-on' tasks motivates students to achieve better results. Further, reading engagement that is initially learned in one knowledge domain can transfer to a new knowledge domain. (Guthrie et al. 1999; Guthrie & Wigfield, 2000, p. 411.) In addition, there is research suggesting that the rich content of a technology environment can encourage 'at-risk' students to participate in literacy experiences and increases both their motivation to become independent readers and writers and their sense of competency (see Kamil et al. 2000). Girls are, on average, more engaged in reading than boys (OECD, 2001, 2002). Boys, however, are more interested in and comfortable with computers

(OECD, 2001; Leino, 2003). Could this be a way to influence boys' reading achievements?

Data and Method

This study focuses on the reading literacy performance of Finnish students and background information provided by the data from PISA 2003. In multilevel modelling of the responses of Finnish students, *confidence in ICT* was the main explanatory variable, while the response variable of the study was the students' reading literacy performance. In order to estimate the response variable, the weighted likelihood estimates were used, instead of plausible values. The weighted likelihood estimates are based on actual responses to reading literacy tasks, but plausible values have also been estimated for those students who did not answer any reading literacy tasks because mathematical literacy was the major domain in 2003. For this study, 2967 students were included in the models. The sample size of schools was 197, and all schools were also present in the reduced student sample[1].

In the models, reading literacy performance was modelled as a function of the variables describing the students' confidence in ICT tasks. Three separate models based on the categories of confidence in ICT (confidence in routine tasks, confidence in Internet tasks and confidence in high-level tasks) were each used as an explanatory variable. In addition, the interaction effects of *gender* and both the linear and quadratic effect were included in the models as well as some important background factors to control for their effects on reading literacy performance. These variables were *language of the school* (Finnish or Swedish), *students' socio-economic background*, *cultural possessions of the family*, and *engagement in reading*. *The students' socio-economic background* was measured using the international socio-economic index of their parents' occupational status (ISEI) (OECD, 2004, p. 307). The PISA index of *cultural possessions of the family* was derived from students' reports on the availability of the classical literature, books of poetry and works of art in their home (OECD, 2004, p. 309). In PISA 2003, questions about engagement in reading, on which the scale of engagement in reading was based, were included in Finland as a national option. *Engagement in reading* was determined as

---

1. To take advantage of the information included in the data structure, at the school level and the student level, and to avoid the problems of intra-class correlated variables, the data were analysed with a multilevel modelling technique (Bryk & Raudenbush, 1992; Goldstein, 2002), using MLwiN software (Rasbash et al., 2000). The statistical method is a two-level regression model, with students as level 1 units and schools as level 2 units.

Table 1 *ICT tasks on three confidence categories and percentages (%) of students' responses*

How well can you do each of these tasks on a computer?	I can do this very well by my self		I can do this with help from someone		I know what this means but I cannot do it		I don't know what this means	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Self-confidence in ICT routine tasks								
Draw pictures using a mouse	94.6	94.4	4.2	4.0	1.1	1.3	.1	.3
Open a file	90.3	96.6	8.1	2.6	1.2	.5	.5	.3
Print a computer document or file	90.0	95.9	6.7	2.6	2.5	.8	.9	.6
Play computer games	89.9	97.5	8.5	1.8	1.2	.5	.3	.3
Save a computer document or file	87.7	95.3	8.2	3.1	3.1	.9	1.0	.7
Delete a computer document or file	85.0	95.3	10.4	3.1	3.6	1.0	1.0	.6
Start a computer game	84.5	96.5	13.1	2.7	2.2	.4	.1	.4
Scroll a document up and down a screen	83.9	91.9	8.7	5.1	3.8	1.9	3.6	1.2
Create/edit a document	66.4	86.3	22.9	10.2	8.1	2.6	2.5	.9
Copy a file from a floppy disk	54.8	90.4	28.6	7.3	13.6	1.6	3.0	.7
Move files from one place to another on a computer	49.9	89.8	26.0	8.0	11.9	1.3	2.2	.9
Self-confidence in ICT Internet tasks								
Get on to the Internet	98.6	97.6	1.2	1.5	.2	.6	0	.2
Write and send emails	95.6	93.2	3.4	4.6	1.0	1.7	.1	.4
Attach a file to an email message	43.6	72.4	35.6	20.1	17.7	5.9	3.2	1.5
Copy or download files from the Internet	44.2	89.5	34.2	7.7	19.6	2.1	1.9	.6
Download music from the Internet	34.7	85.0	34.5	10.1	28.2	4.1	2.5	.8
Self-confidence in ICT high level tasks								
Use a database to produce a list of addresses	31.4	61.5	40.9	27.9	16.5	6.4	11.1	4.2
Use a spreadsheet to plot a graph	28.4	53.6	38.3	33.5	26.1	10.1	7.2	2.8
Create a PowerPoint presentation	28.2	55.8	30.3	28.9	27.1	11.5	14.4	3.9
Construct a web page *	13.5	40.9	44.7	38.3	38.6	18.3	3.2	2.4
Use software to find and get rid of computer viruses	13.4	64.3	31.9	23.1	49.6	11.2	5.1	1.4
Create a multi-media presentation (with sound, pictures, video)	12.5	44.0	33.8	37.9	46.1	15.8	7.6	2.3
Create a computer program (e.g. in Logo, Pascal, Basic)	4.8	24.0	27.4	39.3	49.5	28.8	18.3	7.9

*This item was originally in a field trial added to Internet tasks, but was moved as it appeared to be rather an indicator of confidence in high-level tasks.

in PISA 2000 from nine items measuring the student's attitude towards reading (see OECD, 2001, p. 223).

The degree of confidence in computer use, based on the PISA framework of students' self-reported ability to use computers, was measured for three categories of ICT tasks: routine tasks, Internet tasks, and high-level tasks (OECD, 2005, 306). The questions presented to students are listed in Table 1. Students were asked to make a choice that fitted best with their own views, their options being (1) I can do this very well by myself, (2) I can do this with help from someone, (3) I know what this means but I cannot do it, and (4) I don't know what this means. The answers of Finnish students are presented as percentages in Table 1. It should be noted that responses to the items measuring confidence in routine and Internet tasks were highly skewed, with only a few students expressing a lack of confidence in performing these tasks. This is reflected in the shape of the scales in figures 2 and 3 as a small variation above the mean value.

Correlations between confidence in ICT tasks and frequency of performing these tasks were analysed using the Pearson correlation coefficient. Correlations for eight Internet and high-level tasks, where information both for confidence and frequency was available, were analysed.

Comparison between Nordic countries

Among the Northern countries participating in PISA 2003, Denmark, Finland, Iceland and Sweden gathered information about students' computer use. Among these, the Swedes and the Icelanders were most active in using computers at home and in using the Internet. The Danes were most active in using computers at school. On average, there was not much difference in *confidence in routine tasks* between the Nordic countries. Finnish students did not trust their skills at creating/ editing a document or moving files as much as students in other countries. Finnish and Danish students had significantly lower levels of confidence in their ability to perform internet tasks than Swedish and Icelandic students (except in 'getting on to the Internet'). For high-level tasks, confidence levels varied: the highest confidence levels were reported by Icelandic students for using databases, by the Danes for spreadsheet use, by the Danes and the Swedes for creating multimedia, and by the Danes and the Icelanders for creating web pages. When mean values of confidence in ICT tasks were compared, the Icelanders clearly had the highest confidence levels in high-level tasks, and the Icelanders and the Swedes were most confident at performing routine and Internet tasks. In Finland, Finnish-speaking students were clearly more confidence at routine tasks and high-level tasks than Swedish-speaking students, who, on the other hand, were more confident a Internet tasks. Gender differences related to ICT tasks are presented in figure 1, which clearly shows

higher confidence levels for boys than girls, but also consistency across Nordic countries, especially among boys. In the next section, the results of the multilevel analysis based on the responses of Finnish students are described.

Figure 1 *Means of three categories of confidence in ICT tasks in Nordic countries according to gender*
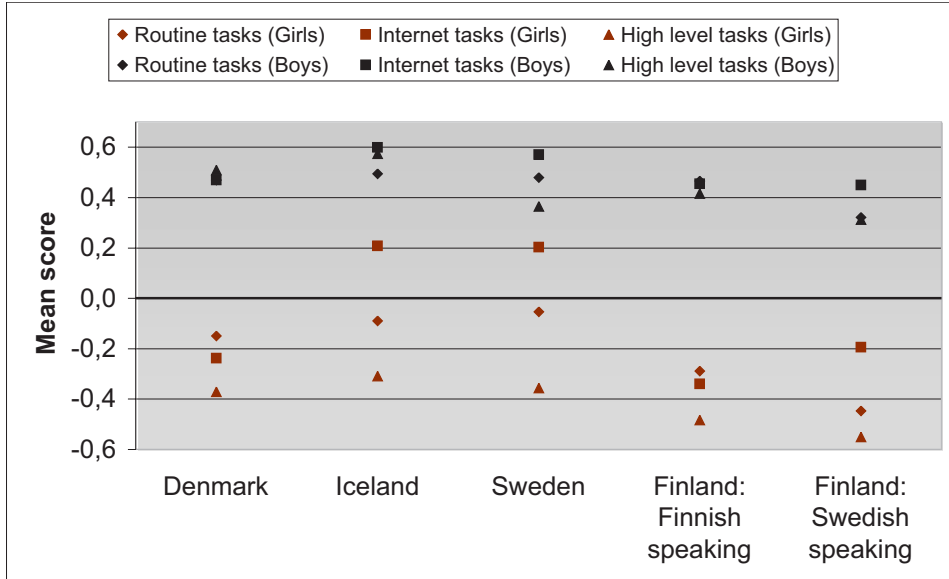

Results

The results of statistical analyses, that is, the association between reading literacy performance and ICT confidence for boys, are presented in Appendix 1[2] to this chapter. There are three models, one for each domain of confidence in ICT tasks. After controlling for the effects of background factors, the linear and quadratic main effects of ICT confidence in reading literacy performance were statistically significant in each model, implying that the association between reading literacy and confidence in ICT tasks was curvilinear. The interaction of gender and linear effect of confidence was statistically significant in all three categories. This means that linear effects are different for boys and girls. The interaction of gender and quadratic effect of confidence was statistically significant in routine tasks only. This means that the shape of the curve, a gently sloping inverted U-shape is different for boys and girls in this ICT domain only.

---

2. For girls, the interaction effects of gender have to be added. The gender differences in Appendix 1 report the difference only when the appropriate confidence factor has the value 0, which was the national mean, and the effects of all other factors are controlled for.

Holding control variables constant, the performance of students with high ICT confidence was better than those with low confidence. However, the effect of computer use was different within routine, Internet and high-level tasks. The results for Internet and high-level tasks are more interesting than those for routine tasks, because the majority of students felt that they could do routine tasks well, even if they used computers only occasionally. In addition, Internet tasks were among the most popular activities involving computers, whereas tasks in the high-level domain require greater technical knowledge and were active hobbies for a smaller group of students (Leino 2002, 2005).

The percentages in Table 1 also reflect the frequencies of different activities; for example, electronic communication, such as email and chat, were the most popular activities among Finnish students (Leino 2002, 2005), which is shown in students perceptions as confidence in Internet tasks. Correlations between students' perceptions of their ICT abilities and reported frequencies of using these particular programs, software or practices were significant. For example, the correlation between perception of ability to download music and frequency of downloading music was 0.67. Correlations ranged from 0.14 to 0.67, with a median of 0.43. In particular, confidence in high-level tasks correlated with frequency of computer use. The most confident students performed these tasks on average at least a few times each week. For Internet tasks, the highest levels of confidence correlated to reported frequencies of task performance of at least a once a month.

Figure 2 *Association between self-confidence in ICT routine tasks and reading literacy score, after controlling for the effects of background factors*
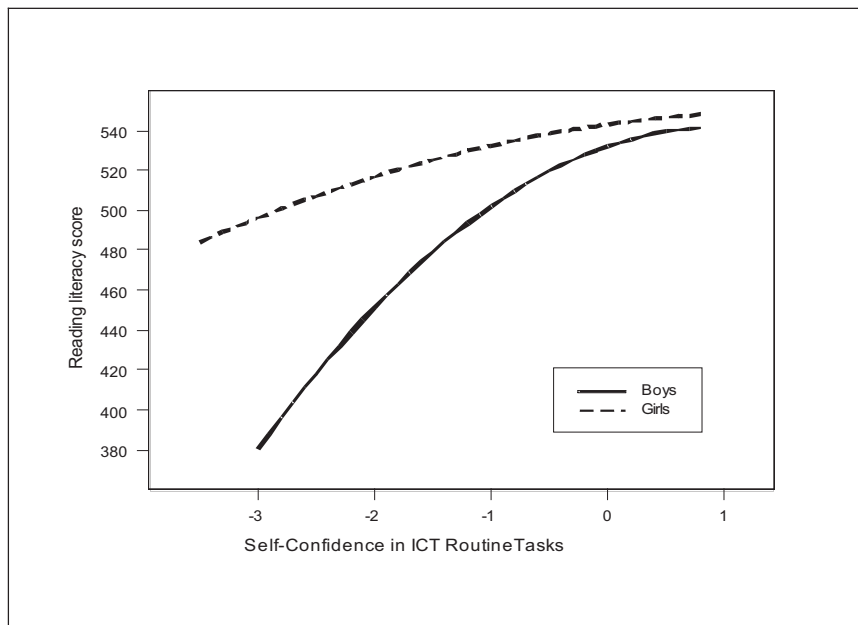
In figure 2, the relationship between confidence in ICT routine tasks and reading scores is presented separately for boys and girls, as the effects of background factors were controlled for. The gender difference depends on the level of confidence in ICT tasks. Finnish students had very high levels of confidence in their routine task skills. Confidence in their skills was clearly positively associated with reading scores, as seen in figure 2. Confidence in routine tasks was more positively associated with boys' reading scores than girls'. For girls, the association was almost linear, while for boys it was clearly curvilinear. Among students with the lowest confidence in routine tasks (-3.0), the gender difference in reading scores was 116 points higher for girls. In terms of proficiency levels that is more than one and a half times higher. But among those with the highest confidence in routine tasks (0.8) the difference was only seven points greater for the girls.

Figure 3 *Association between self-confidence in ICT Internet tasks and reading literacy score, after controlling for the effects of background factors*
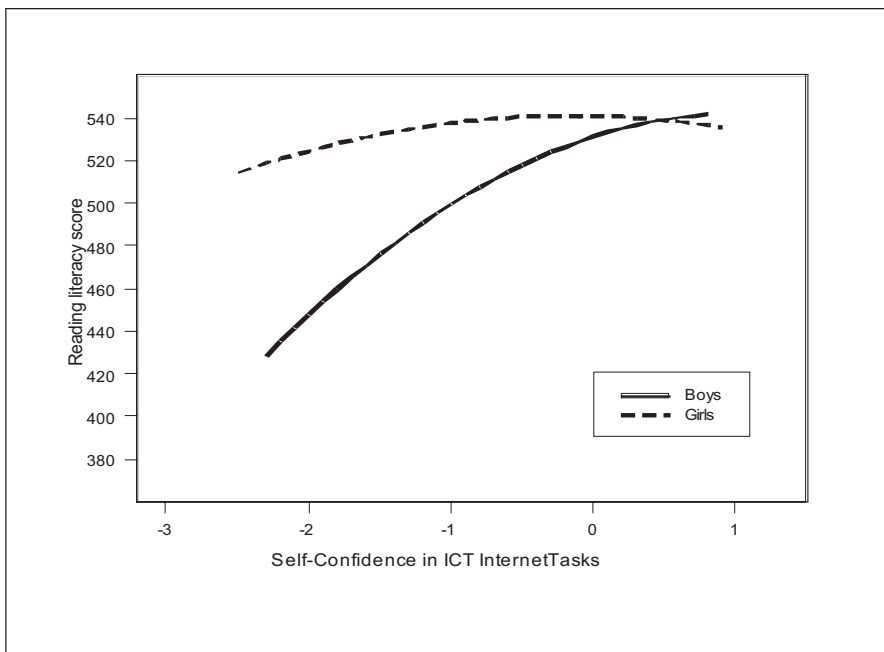
Students' confidence in their skills in Internet tasks had almost the same kind of relationship to reading scores as with routine tasks, as seen in figure 3. When controlling for the effects of background factors, the relationship was stronger for boys: Among those with the lowest confidence in their Internet skills (-2.0) the gender difference was 77 points in favour of girls but among the most confident students (0.8) the difference was five points higher for boys.

Figure 4 *Association between self-confidence in ICT high level tasks and reading literacy score, after controlling for the effects of background factors*
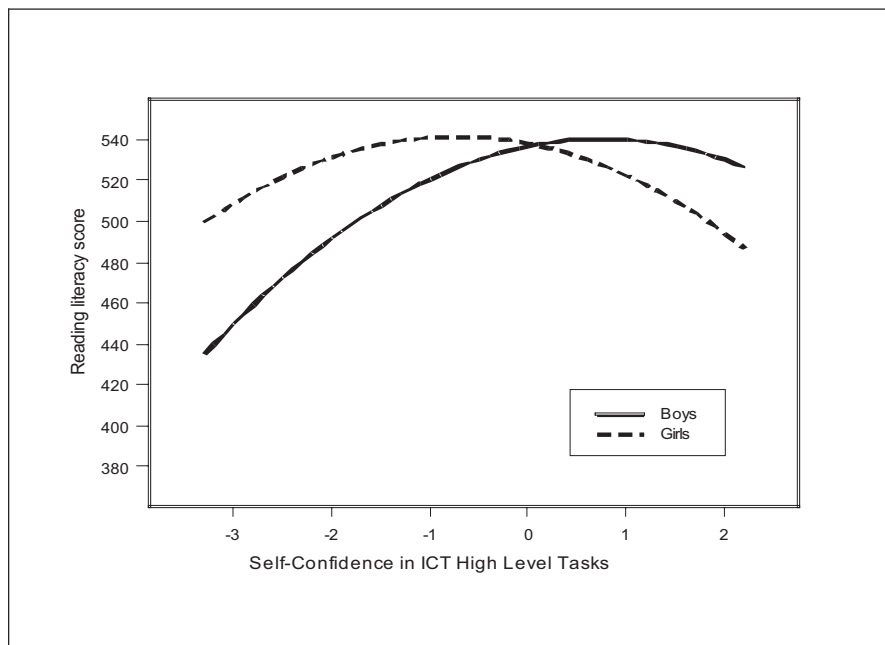
The relationship between high-level tasks and reading scores was a gently sloping inverted U-shaped curve, with student achievement increasing and decreasing with the level of confidence, as seen in figure 4. An interesting fact to note is that the lines for boys and girls crossed. Among students with clearly below-average confidence in their skills in high-level tasks, i.e. those with a confidence value of -3, the gender difference was 59 points higher for girls. But among the most confident students, i.e. those with a value of 2, the difference was 36 points higher for boys. However, the best reading scores were achieved by boys with confidence levels a little above average and by girls with confidence levels a little below average.

In addition to the effects of variables describing confidence in ICT tasks, the coefficient estimates of background factors are also presented in Appendix 1. As would be expected, all three models showed equal results: *higher socio-economic background*, *more cultural possessions in the family* and *more engagement in reading* were all associated with better results in reading literacy performance, on average. There was no statistically significant difference caused by *language of the school*. That is interesting, because in general Swedish-speaking students did not perform quite as well in reading literacy as Finnish-speaking students, the difference being small but statistically significant (Linnakylä & Sulkunen, 2005). Due to the interaction between them, the estimates of gender difference in the models have to

be combined with confidence in ICT tasks, on which the gender difference depends.

The background factors and the ICT confidence variables together explained 22-24% of the total student variance in reading literacy performance. However, the background factors alone explained about 20% of the variance. The between-school variation in reading literacy performance was small in Finland. The unadjusted intra-class correlation in this reduced sample indicated that only 3.3% of the total student variance in reading literacy performance was attributable to the differences between schools. The between-school variance component was reduced by 44-63% of the already originally small between-school variation. It was notable that confidence in routine tasks reduced the between-school variation most. However, small school differences existed after controlling for the effects of both confidence in ICT tasks and background factors.

Discussion

Girls, on average, have less confidence in their ICT skills than boys. According to our results, boys seem to derive advantage from using computers which is reflected in a high level of confidence in their ICT skills. On average, even though the boys with most confidence in their high-level ICT skills did not do as well in reading as boys with moderate confidence, they still did better than girls. The results show that high self-confidence in ICT has a positive relationship with reading achievement. Confidence in the most difficult tasks, high-level tasks, are especially reinforcing for boys. This result differs from earlier study, in which the relationship between individual high-level tasks and reading proficiency in PISA was studied without controlling other variables (see Leino, 2002), as was done here. On the other hand, the present results confirm other results based on PISA 2000 showing that Internet use and moderate computer use are the most advantageous for reading achievements (Fuchs & Woessmann, 2004; Leino, 2002; Leino et al., 2004).

One factor explaining the results may be the literacy practices of students who use computers: Active users of high-level tasks are also active readers of comics and non-fiction (Leino, 2002), kinds of texts that are also widely found on the Internet. In particular those students who are eager to learn more about areas such as programming or constructing a web site actively read related texts such as printed manuals or related discussions on the Internet. Those students most confident in high-level tasks are probably the heaviest users of ICT, who spend their time entirely on computers with no interest in other literacy practices. It is self-evident that, for example, one-sided program writing does not strengthen high-level reading proficiency, which demands skills and knowledge to critically evaluate, understand nuances of language, analyse and match information, and fully understand long texts (OECD, 2002, p. 40). These skills can be achieved through

use of diversified reading materials and engagement in reading (OECD, 2002). The results of a Finnish students' reader profile study (Leino et al. 2004) also support this interpretation: namely that the heavy users of the Internet were not as diversified readers, in terms of media texts and, especially, traditional texts (fiction and non-fiction), as those of somewhat more moderate users.

Proficiency in reading literacy is associated with many factors (e.g. reading engagement, social background, and diversity of reading materials). Even though confidence in computer use and related activities can explain only a few per cent of variance in reading literacy, it is still an important factor, because it is something we can quite easily control. In this study students' ability to use computers was assessed from their self-reported confidence levels. This study does not exclude the possibility that one effective variable is actually the student's confidence in his/her own abilities in general, not just self-reported confidence in using computers. Indeed, this relationship can be two-way street also, as is so often the case where reading is concerned. Other results from PISA 2000 showed that high self-confidence was related to good performance. However, on average, boys' self-confidence related to reading was much lower than girls'. Reinforcing the ICT skills of students with low levels of confidence in ICT and encouraging those skills in school may be one factor in equalize gender differences in reading.

As cognition, motivation, proficiency and engagement in reading have an entangled relationship (OECD, 2002), access to interesting and meaningful reading materials is important. Electronic texts could clearly be one way to increase boys' interest in reading and their proficiency levels as well. By developing students' ICT skills and practices we can also develop opportunities and access beyond formal education and increase students' motivation for learning. Reading outside the school environment can be motivating, and any kind of reading is better than nothing. Computers and the Internet are still very much text based. Old and new learning environments can and must be used in a complementary fashion to promote multiliteracy practices which are a significant factor in the modern media world. Indeed, researchers have even found a pattern that suggest that students 'at risk' in school literacy are "sometimes the most adept at (and interested in) under-standing how media texts work" (Alvermann, 2002, p. 17).

In the PISA survey only print texts were assessed which obviously omits the assessment of features characteristic of ICT literacy. If the assessment had included electronic texts, we can assume that the relationship between confidence in ICT and reading literacy would have been stronger. However, these results indicate that use of ICT and positive confidence in ICT may also reinforce reading print literacy. We await with interest future literacy assessments which will hopefully give us more information about ICT literacy skills of young people.

Appendix 1 *The effects of confidence in ICT tasks on reading literacy performance, after controlling for the effects of background factors*

Fixed effects	Confidence in routine tasks			Confidence in Internet tasks			Confidence in high level tasks		
	b	s.e.	p	b	s.e.	p	b	s.e.	p
Intercept*	532.4	4.16	0.000	531.5	3.86	0.000	537.3	2.64	0.000
Confidence in ICT tasks:									
Linear effect	19.7	3.26	0.000	21.3	2.90	0.000	9.6	2.46	0.000
Quadratic effect	-10.3	3.37	0.002	-10.2	4.05	0.012	-6.5	1.49	0.000
Gender by Linear effect	-11.6	4.24	0.006	-22.9	4.08	0.000	-19.0	4.01	0.000
Gender by Quadratic effect	7.8	3.94	0.047	5.3	5.03	0.290	0.1	2.32	0.965
Controlling background factors:									
Gender (girls)	10.9	4.96	0.027	9.7	4.80	0.044	1.2	3.67	0.735
Language of school (Swedish)	-9.5	5.94	0.110	-11.2	6.08	0.064	-10.9	6.12	0.076
Socio-economic index	0.7	0.09	0.000	0.7	0.09	0.000	0.8	0.09	0.000
Cultural possessions	4.5	1.42	0.002	5.3	1.43	0.000	5.8	1.44	0.000
Engagement in reading	27.1	1.62	0.000	28.1	1.64	0.000	27.9	1.65	0.000
Random effects									
Between-school variance component	86.1	42.3		122.7	46.8		131.4	48.3	
Within-school variance component	5222.7	139.5		5292.9	141.4		5345.0	142.8	
Total variance	5308.8			5415.6			5476.4		
ICC	0.016			0.023			0.024		
Reductions in variance components (%)									
Between-school variance component	63.1			47.5			43.8		
Within-school variance component	22.7			21.7			20.9		
Total variance	24.1			22.5			21.7		
N of students	2967			2967			2967		
N of schools	197			197			197		

* Intercept is the expected reading literacy score for students whose all background factors are equal to 0.

References

Alvermann, D. E. (2002). *Effective Literacy Instruction for Adolescents*. Executive Summary and Paper Commissioned by the National Reading Conference. Retrieved August 25, 2005, from http://www.coe.uga.edu/reading/faculty/alvermann/effective.pdf

Bandura, A. (1986a). *Social Foundations of Thought and Action. A Social Cognitive Theory*. New Jersey: Prentice Hall.

Bandura, A. (1986b). *Definition of self-efficacy*. Retrieved June 8, 2005, from http://www.emory.edu/EDUCATION/mfp/efficacy.html

Bryk, A., & Raudenbush, S. (1992). *Hierarchical Linear Models: Application and Data Analysis Methods*. Newbury Park: Sage Publications.

Fuchs, T. & Woessmann, L. (2004). *Computers and student learning: bivariate and multivariate evidence on the availability and use of computers at home and at school*. CESifo working paper no. 1321. [Electronic version] Retrieved May 29, 2005, from http://ideas.repec.org/p/ces/ceswps/_1321.html

Goldstein, H. (2002). *Multilevel Statistical Models*. Third Edition. London: Edward Arnold.

Guthrie, J. T., Anderson, E., Alao, S. & Rinehart, J. (1999). Influences of Concept-Oriented Reading Instruction on strategy use and conceptual learning from text. *Elementary School Journal 99 (4)*, 343-366.

Guthrie, J. T. & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.) *Handbook of reading research (Vol. 3)*. Mahwah, NJ: Erlbaum. Pp. 403-422.

Kamil, M. L., Intrator, S. M. & Kim H. S. (2000). The effects of other technologies on literacy and literacy learning. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.) *Handbook of reading research (Vol. 3)*. Mahwah, NJ: Erlbaum. Pp. 771-788.

Leino, K. (2002). Tietotekniikan käyttö ja lukutaito. In J. Välijärvi & P. Linnakylä (Eds.) *Tulevaisuuden osaajat - PISA 2000 Suomessa*. Jyväskylä: Institute for Educational Research. Pp. 167-180.

Leino, K. (2003). Computer usage and reading literacy in the Nordic countries. In S. Lie, P. Linnakylä & A. Roe (Eds.) *Northern Lights on PISA. Unity and diversity in the Nordic Countries in PISA 2000*. University of Oslo. Department of Teacher Education and School Development.

Leino, K. (in press). Reading the Web – Students' Perceptions about the Internet. *Scandinavian Journal of Educational Research*.

Leino, K., Linnakylä, P. & Malin, A. (2004). Finnish students' multiliteracy profiles. *Scandinavian Journal of Educational Research 48 (3)*, 251-270.

Linnakylä, P. & Sulkunen, S. 2005. Suomalaisnuorten lukutaito ja -harrastus. In P. Kupari & J. Välijärvi (Eds.) *Osaaminen kestävällä pohjalla. PISA 2003 Suomessa*. Jyväskylä: Institute for Educational Research. Pp. 37-64.

Pajares, F. (2002). *Self-efficacy beliefs in academic contexts: An outline*. Retrieved August 23, 2005, from http://www.emory.edu/EDUCATION/mfp/efftalk.html

Pajares, F. & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.) *Perception*. London: Ablex Publishing. Pp. 239-266.

OECD (2001). *Knowledge and Skills for Life. First results from PISA 2000*. Paris: OECD.

OECD (2002). *Reading for Change. Performance and engagement across countries. Results from PISA 2000.* Paris: OECD.

OECD (2004). *Learning for Tomorrows World. First Results from PISA 2003*. Paris: OECD.

OECD (2005). *PISA 2003 Technical Report*. Paris: OECD.

OECD (2006). *Are Students Ready for a Technology-Rich World? What PISA Studies Tell Us*. Paris: OECD.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I. & Lewis, T. (2000). *A user's guide to MLwiN. Version 2.1b.* Multilevel Models Project. Institute of Education, University of London.

Wenglinsky, H. (1998). *Does it compute? The relationship between educational technology and student achievement in mathematics*. Princeton, NJ: ETS Policy Information Center. [Electronic version]. Retrieved August 25, 2005, from ftp://ftp.ets.org/pub/res/technolog.pdf

Chapter 13

Scientific Competence and Educational Reforms in Norway and Sweden

Karl Göran Karlsson, Marit Kjærnsli, Svein Lie og Maria Åström

Abstract

In this article we present and discuss changes in the science competencies of Norwegian and Swedish students from 1995 to 2003, based on analyses of data from PISA and TIMSS. Both Sweden and Norway implemented major educational reforms in 1994 and 1997, respectively. PISA and TIMSS therefore contain relevant data that can be used for evaluating these reforms. The TIMSS data show a distinct drop in the students' performance in science. This also holds true for the Norwegian students' performance in PISA, but the decline is less significant for Swedish students. The extent to which these changes can be explained by the reforms is discussed. Differences in changes in performance between girls and boys are also considered, as are differences between high- and low-performing students. In some respects the two datasets show opposite tendencies, which can probably be ascribed to the fact that different competencies are measured in PISA and TIMSS.

Nordic abstract

I denne artikkelen vil vi presentere og diskutere endringer i norske og svenske elevers naturfaglige kompetanser fra 1995 til 2003, basert på analyser av data fra PISA og TIMSS. Både Sverige og Norge hadde store skolereformer i henholdsvis 1994 og 1997. PISA og TIMSS har derfor gode data som gir mulighet for vurdering av disse reformene. I TIMSS er det en klar tilbakegang i elevenes naturfagprestasjoner, for de norske elevene gjelder dette også i PISA mens tendensen er svakere hos de svenske elevene. I hvilken grad disse endringene kan tilskrives reformene blir diskutert i artikkelen. Videre blir forskjeller mellom jentenes og guttenes utvikling i prestasjoner tatt opp. Det samme gjelder utviklingen til gruppene av høyt- eller lavtpresterende elever. På noen områder peker ikke resultatene fra de to undersøkelsene i samme retning, noe som kan relateres til at det er ulike kompetanser som måles i PISA og TIMSS.

Introduction

In December 2004, when results from PISA 2003 were presented, much focus in media was on ranking lists and overall changes since PISA 2000. In this chapter we will discuss some further findings concerning Norwegian and Swedish data, in particular in the domain of science literacy. In Sweden, no significant changes between 2000 and 2003 occurred in any of the three test domains in PISA (reading literacy, mathematical literacy and scientific literacy). Moreover, Swedish scores were significantly above the OECD average in all three domains. In Norway, there were no significant changes in results in reading and mathematics, but the drop in science scores was highly significant. The mean Norwegian science score was also distinctly below the OECD average, whereas the reading and mathematics scores were close to the international means (OECD 2001, 2004). Strictly speaking, PISA measures what is called reading, mathematical and scientific *literacy* (for definitions and discussion see e.g. OECD, 2003; Olsen, 2005), but in this chapter we use the more convenient terms reading, mathematics and science.

We will take a closer look at the PISA science results for the two countries in order to discuss possible links to educational reforms that have been introduced during the last decade. For that purpose we will also discuss some results from TIMSS (Trends in International Mathematics and Science Study), or more precisely, results for the science 'grade 8', part of the study (Martin et al. 2004). Besides being discussed in the national PISA 2003 reports for Sweden (Skolverket, 2004a) and Norway (Kjærnsli et al. 2004), the drop in achievement in science is also discussed in the national TIMSS 2003 reports (Skolverket, 2004b; Grønmo et al. 2004). Analyses of Kjærnsli et al. (2005) have recently focused quite specifically on the general decline in competencies in the Nordic countries during the last decade.

PISA science results in more detail

Science was a minor domain in both PISA 2000 and PISA 2003. Results for science were presented on the same scale in both assessments and are thus directly comparable. In each of the two studies there were 34 items in the science test, of which 25 were used as link items to ensure comparability. Due to the small number of items, only general science (or rather scientific literacy) results have been calculated and reliable analyses of the results for the sub-domains of science were not possible. This will, however, be possible in the PISA 2006 assessment, when science is the particular focus.

As stated above, science results showed a significant decline from 2000 to 2003 in Norway. In fact, of all the countries participating in PISA, Norway showed one of the largest achievement drops (after Austria and Mexico). There was also a drop

in the mean science score in Sweden, but this was not quite statistically significant. A closer look at the results reveals that there are other important features of the science results in both countries. Figure 1 displays the changes in the mean scores for the two countries, as well as for each gender separately. From the figure it can be seen that the decline is clearly larger in Norway. The figure also clearly demonstrates that the girls' scores in both countries have decreased more than the boys'.

Figure 1 *Mean scores from PISA 2000 and PISA 2003 for all students and by gender. The score values are defined by setting the OECD mean to 500 and the standard deviation to 100 score points*
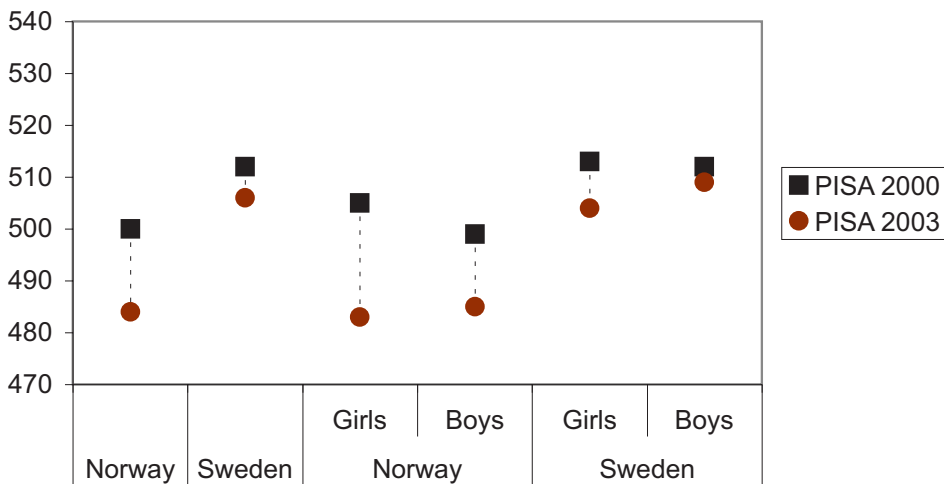
Let us now investigate the performance drop in more detail by studying how the scores have changed for different ability groups. The standard deviations increased from 93 in PISA 2000 to 107 score points in PISA 2003 for Sweden and from 96 to 104 score points for Norway. Combined with the overall decline, this could indicate that the drop is particularly pronounced for the low-ability group of students. Figure 2 displays changes in the score points of students at different percentiles. For both countries, but particularly for Sweden, the decline is much stronger for the weaker students. However, the overall decline is larger in Norway, as discussed above.

Although the changes in Norway are obviously quite dramatic, the Swedish science results also give cause for concern. A cornerstone of Nordic educational policy is that all students should reach an acceptable knowledge level, but the changes revealed here clearly show a worrying trend towards lower standards.

Figure 2 *Changes in science score points from PISA 2000 to PISA 2003 for different points of the score distribution*
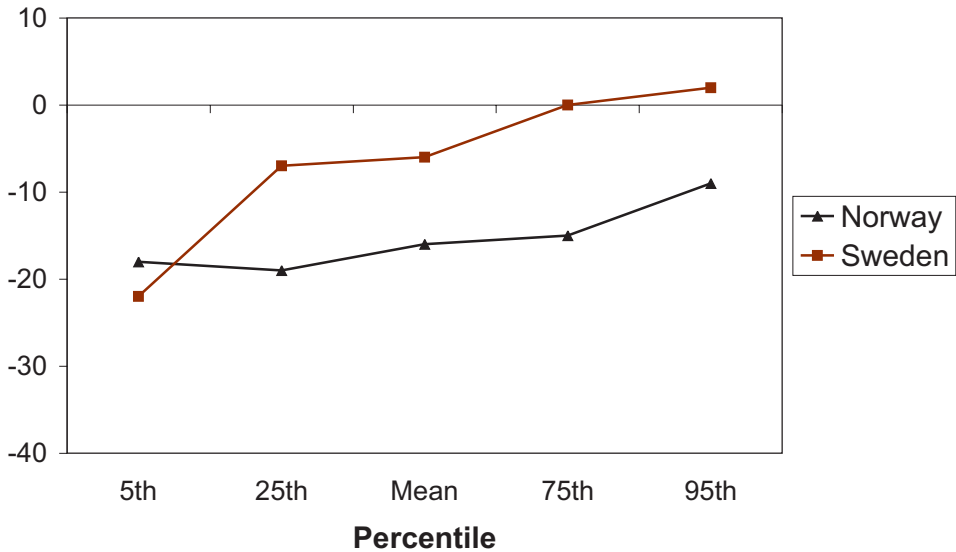
It could be suggested that decline in the science results for Norway and Sweden merely reflect general changes unrelated to changes in students' competence. For instance, some items in the 2000 test were replaced in 2003 by new items, and one could argue that the new items might have had different characteristics. However, before these items were accepted for inclusion in PISA 2003, they were checked to ensure that they could be assessed on the same science scale as the items in 2000 for all participating countries. A closer inspection of both sets of results clearly reveals that the science results, especially at the low performance end of the scale, have dropped not only in absolute terms, but also compared to results for other participating countries.

Comparison with TIMSS data

PISA and TIMSS are different in many respects (e.g. Olsen, 2005, pp. 23-32). The populations and samples are different: TIMSS tests grade 8 students (in most countries, but there were substantial age differences, e.g. between the two countries discussed here) and samples entire classrooms, and while PISA measures what is termed scientific literacy, which is rather loosely tied to school content, TIMSS is much more driven by the curricula of the participating countries. Still, since both studies measure performance of some science competencies, a comparison with TIMSS is meaningful.

There were 18 countries participating in both TIMSS 1995 and TIMSS 2003, among them Sweden and Norway, where age groups were comparable in the two studies. However, direct comparison between the two countries is not very meaningful because the age groups are different – Swedish TIMSS students are about one year older than their Norwegian peers. However, intra-national comparisons between 1995 and 2003 results can be done for each of the 18 countries (Martin et al. 2004). Figure 3 displays the change in scores for these countries. The downward trends for the two countries are immediately striking, since they occupy the two lowest places on the diagram.

Figure 3 *Changes in science score from TIMSS 1995 to TIMSS 2003 for countries participating in both studies*
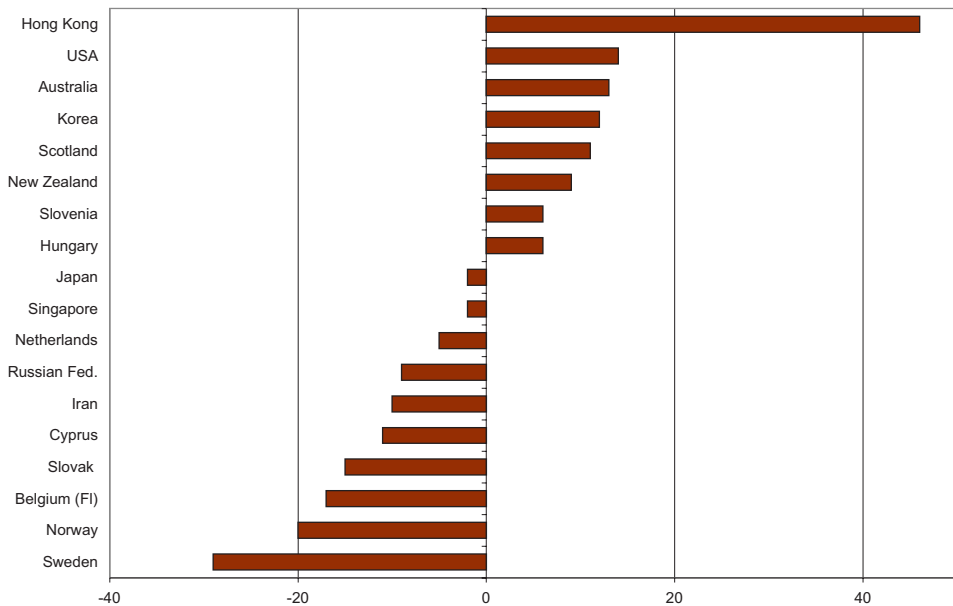
More detailed data on trends are shown in figures 4 and 5. Figure 4 specifies the performance drop by gender, whereas figure 5 compares the performance drop for students at different percentiles along the science scale.

Together, figures 4 and 5 call for comment, since in many respects they contradict the findings from PISA. Firstly, the bigger achievement drops are associated with boys instead of girls, and with Sweden instead of Norway. Secondly, the trends for high- and low-performing students are clearly reversed, since the drop is particularly strong for the best students, while the profiles for the two countries are essentially the same.

Figure 4 *Changes in mean scores from TIMSS 1995 to TIMSS 2003 for all students and by gender*
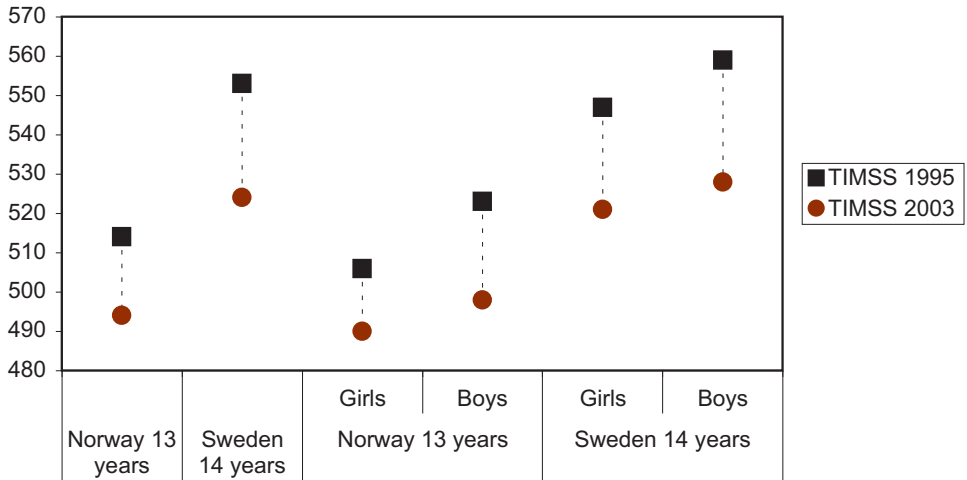
Figure 5 *Changes in science score points from TIMSS 1995 to TIMSS 2003 for various points at the score distribution*
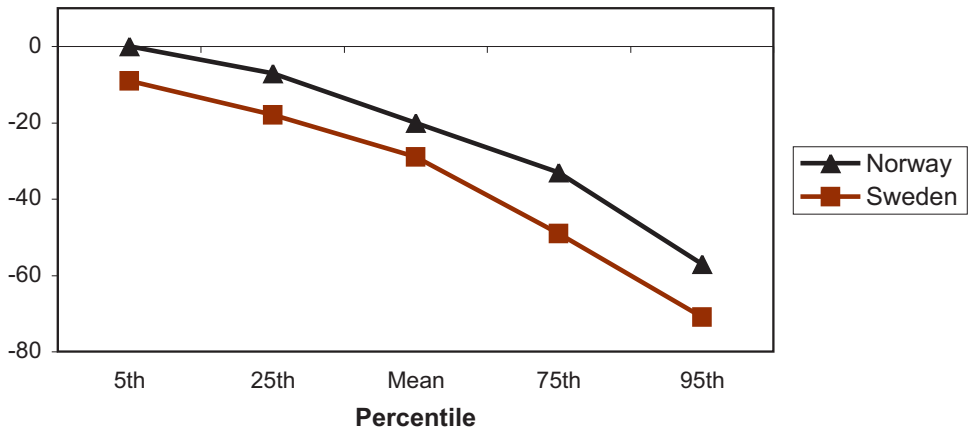
Discussion

Conflicting evidence from PISA and TIMSS?

The contrasting pictures obtained from TIMSS and PISA should be interpreted in the light of the important differences in the time intervals between the two studies and the other differences discussed above. Most important to consider here is how the timing of the studies relates to that of educational reforms, and also how the content tested in the two studies relates to the national curricula. Major educational reforms for compulsory schooling were implemented in Sweden and Norway in 1994 and 1997, respectively. Therefore, the changes seen in TIMSS from 1995 to 2003 may be regarded partly as a consequence of these reforms. In contrast, there are no PISA data from the years before the reforms were implemented.

In the following section we will discuss the findings from each country's perspective. But first, one important aspect should be specifically mentioned, because it is very relevant to the interpretation of the different profiles of decline in figures 2 and 5. It must be remembered that TIMSS is closely linked to some kind of 'average' curriculum, and the test is strongly based on content coverage. PISA, on the other hand, has a much stronger emphasis on what are often called 'process items', i.e. items that largely depend on scientific reasoning. These items do not depend on detailed content knowledge. For bright students this creates a big difference between the two studies. If the reforms in Sweden and Norway have in practice resulted in a less detailed focus on more advanced content, this will mean that the high-performing students will have been less well prepared by their school teachers for the most demanding items in TIMSS than in PISA. Consequently, this difference offers a potential explanation of the difference between the two profiles. Even the brightest students will struggle with TIMSS items that strongly depend on particular content knowledge or conceptual understanding that has not yet been taught.

Sweden

In Sweden a new curriculum was introduced in 1994. The reform was characterised by more emphasis on process skills like reasoning, arguing and working together than on specific factual knowledge. As part of the reform, a goal-oriented system, with specific criteria for a pass grade in each subject, was introduced. Syllabuses for different subjects were again revised in 2000. Moreover, there was a grading reform in 1998, when it was decided that students must achieve pass grades in Swedish, English and Mathematics to enter a national programme in upper secondary school (SOU 1997:121, p. 161). This system was used for the first time in the final year of compulsory school in 1999.

Can the Swedish results in PISA and TIMSS in any way be related to the reforms in the 1990s and 2000? Because of the limited number of measurements

available, it is hard to draw any definite conclusions, but the results in the assessments do call for a discussion.

There are observations from other assessments that could explain the decline in science results seen in TIMSS. In a national assessment, also conducted during spring 2003, a significant decrease in students' conceptual knowledge in science was found (Skolverket, 2004c). On the other hand, the same investigation also demonstrated an increase in process skills such as arguing, which was the intention of the curricular reform in 1994. In the Swedish TIMSS report (Skolverket, 2004b) it is pointed out that a thorough analysis of the items and their relevance to the national curriculum needs to be done in order to be able draw any certain conclusions about causes for the observed decline. Lacking such an analysis, we can only speculate that students have realised that features of the curriculum other than conceptual knowledge are more rewarding in getting good grades. Molander (1997) has shown that high performers in particular are good 'cue-seekers', meaning that they are good at identifying where effort pays off. If this is the explanation it would not necessarily mean that students learn less, but that they may learn more about other skills than those being tested.

There is, however, a more worrying interpretation. In the mid-1990s, just after the curricular reform, there was considerable focus on changes in teaching styles. The debate was particularly intense in the science subjects, because it had been noticed that many young people lost interest in these subjects towards the end of the compulsory school period. It was argued that there should be less lecturing and more student-centered activities like group work and projects, and in experimental work students should 'find things out for themselves' without too much interference and guidance from the teacher. In short, teaching should be less formal. However, it has long been argued that most students, and especially students of low ability, do not benefit from this kind of informal teaching (e.g. Bennett, 1976). Bergqvist (1999) has also written a very critical report of this 'exploratory pedagogy'. If changes in teaching styles lie behind the drop in students' test scores it is likely that students really have learned less science than they did before.

The decline in the PISA results is equally difficult to interpret. The poorer performance of the weaker students has already been discussed in the Swedish national PISA report (Skolverket, 2004a). At that time there was not much additional support for the idea presented in that report - that the decline could be due to the increased emphasis put on the 'core subjects' Swedish, English and Mathematics. In a recent study Eriksson et al. (2004) discuss the effects of the policy that every student wanting to enter a national programme in upper secondary school must have a pass grade in those three subjects. These authors have interviewed a substantial number of teachers to evaluate the effects of a five-year pilot project for working with no set timetable. Some quotes from their report illustrate what happens (all are our translations):

"Then there are imperfections in the system since some subjects are valued more highly or need inevitably to be passed to get into the upper secondary school. And then you cannot let the students make their own choice." (p. 41)

"And then maybe this requirement of eligibility in only three subjects to get into upper secondary school, it is also a risk. You need to get passed in maths, so then we take something else away and push in more maths for example." (p. 41)

The authors of the report conclude:

"Teachers interpret their task as guaranteeing a three-subject school, where their mission is to make sure that students get at least a pass grade in Swedish, English and Mathematics." (ibid p. 43f; our translation)

We feel that the report strongly supports the idea put forward in the national PISA report (Skolverket, 2004a), that less time and effort is put into other subjects than into the three 'core subjects', and that low-performing students suffer from this. The lowest achievers show a slight increase in mathematics scores from PISA 2000 to PISA 2003. More interesting, however, is the fact that the proportion of students not reaching the goal of a pass in the national mathematics test was much smaller in 2003 than in 2000. This supports the idea that a high priority is given to mathematics.

Norway

The educational reform of 1997 in Norway was implemented gradually, starting in the school year 1997-98 (for grades 1, 2, 5 and 8). Two years later all students were 'reform students'. In particular, the PISA students (mainly grade 10) had followed the reformed curriculum for the full three-year lower secondary period ('ungdoms-skole', grades 8-10). Curricula for each subject and each grade were described mainly in terms of detailed instructional tasks and subject matter content. The learning goals for students were, however, rather vague. Therefore, there was a tendency to blame the reforms when the PISA 2000 results, with worse-than-expected mean achievement scores in all domains, were reported. By the time the result were reported, the education minister had already announced work towards a new curricular reform ("Kunnskapsløftet"). One of the proposed cornerstones of this reform (later implemented) was the idea of freedom of instructional methods combined with concrete descriptions of *what* content knowledge and skills students are expected to learn and therefore to be able to use. Focus should be on learning *goals* rather than on learning *activities*. In addition, some skills were given specific emphasis as so-called "basic skills" due to their crucial role in further learning. Detailed plans for national assessments in these areas were also announced (and implemented in 2004). The specified skills included basic reading, writing, English, mathematics and ICT skills, but science competencies were not considered

sufficiently 'basic' to be included. Thus, the lower priority given to science compared to other subjects in Norway clearly parallels the situation in Sweden.

The PISA 2003 results were similar to those of 2000 in maths and reading, but the decline in science (as well as the surprisingly low scores in general problem solving) led to further concern. This decline in science achievement could, however, partly be linked to the decreased emphasis on science that has already been signalled (see above) before the reform is actually due to be implemented (in 2007).

The decline in science results in TIMSS may also be related to the reforms of 1997, more so than for PISA, since the 1995 TIMSS data reflect the situation prior to the implementation of the reforms. In fact, TIMSS 2003 was specifically announced by the ministry as a tool to measure the effect of the reform. And the outcome, the general decline in both science and maths and in results for both grades 8 and 4, defined much of the official argument for claiming the 1997 reform to be a failure and in need of revision.

In the national reports for PISA 2003 (Kjærnsli et al. 2004) and TIMSS 2003 (Grønmo et al. 2004) the pedagogical situation was discussed thoroughly. In particular, other types of data from the two studies were also considered. These included data illustrating particular problems in Norway concerning the disciplinary climate in the classroom, teachers' backgrounds in subject matter, teacher-student relations and students' attitudes towards schooling and subjects. The general 'modern' trend away from teacher-led instruction towards student-led activities was also questioned and has since been in the focus of much scholarly and media debate.

Conclusion

In both countries the science results from PISA and TIMSS have provided important evidence for a substantial decline in lower secondary students' scientific competencies during the last decade. This issue is of great concern, and some action has been taken to counteract this tendency. The next phase of the PISA and TIMSS studies may provide some evidence of whether these actions have had any effect. For the PISA 2006 study, when science will be the main subject domain, subcategories for science w ill also be included. The larger number of items will yield a more solid foundation for understanding what may have gone wrong.

References

Bennett, N. (1976). *Teaching styles and Pupil Progress*. London, Open Books.

Bergqvist, K. (1999). Vi lärde oss bara sätta fast lampor och grejor på bänken. *Pedagogiska magasinet nr 3/99, pp. 27-31.*

Grønmo, L.S., Bergem, O.K., Kjærnsli, M., Lie, S. & Turmo, A. (2004). *Hva i all verden har skjedd i realfagene? Norske elevers prestasjoner i matematikk og naturfag i TIMSS 2003.* IILS, University of Oslo

Eriksson, I., Orlander, A.A., Jedemark, M. (2004). *Att arbeta för godkänt – timplanens roll i ett förändrat uppdrag*. Centrum för studier av skolans kunskapsinnehåll i praktiken. Stockholm, HLS Förlag.

Kjærnsli, M., Lie, S., Olsen, R.V., Roe, A. & Turmo, A. (2004). *Rett spor eller ville veier? Norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003.* Oslo: Universitetsforlaget.

Kjærnsli, M., Lie, S. & Turmo, A. (2005). Kan elevene mindre enn før? Naturfagkompetanse i Norden i perioden 1995-2003. *Nordina* 1 (2) p. 51-60.

Martin, M. O., Mullis, I.V.S., Gonzales, E.J., Chrostowski, S.J. (2004). *TIMSS 2003 International Science Report.* TIMSS & PIRLS International Study Center, Boston College

Molander, B-O (1997). *Joint discourses or disjointed courses.* Studies in Educational Sciences 8. Stockholm, HLS Förlag.

OECD (2001). *Knowledge and Skills for Life. First Results from PISA 2000.* Paris: OECD.

OECD (2003). *The PISA Assessment Framework. Mathematics, Reading, Science and Problem solving Knowledge and Skills*. Paris: OECD.

OECD (2004). *Learning for Tomorrow's World. First Results from PISA 2003*. Paris: OECD.

Olsen, R.V. (2005). *Achievement tests from an item perspective. An exploration of single item data from the PISA and TIMSS studies, and how such data can inform us about students' knowledge and thinking in science.* Dr. thesis, Faculty of Education, University of Oslo

Skolverket (2004a). *PISA 2003 – Svenska femtonåringars kunskaper och attityder i ett internationellt perspektiv.* Rapport 254. Stockholm: Skolverket.

Skolverket (2004b). *TIMSS 2003. Svenska elevers kunskaper i matematik och naturvetenskap i skolår 8 i ett nationellt och internationellt perspektiv* Rapport 255. Stockholm: Skolverket.

Skolverket (2004c). *Nationella utvärderingen – sammanfattande huvudrapport.* www.skolverket.se

SOU 1997:121 (1997). *Skolfrågor - Om skola i en ny tid*. Utbildningsdepartementet. Stockholm

Chapter 14

Gender and the Urban-rural Differences in Mathematics and Reading: An Overview of PISA 2003 Results in Iceland

Ragnar F. Ólafsson, Almar M. Halldórsson and Júlíus K. Björnsson

Abstract

Gender and urban-rural differences in the Icelandic PISA results are examined in this chapter, and we show that the unusual Icelandic gender difference, where girls outperform boys in mathematics, is a complex phenomenon not readily explained by any current popular theory or understanding of gender differences. A comparison of the PISA 2003 results with the Icelandic National Examinations scores for the same students confirms the validity of the PISA result, but also clearly reveals that gender differences in mathematics in Iceland fluctuate considerably from year to year and from region to region, showing that the difference is very inconsistent and therefore difficult to understand. The results also indicate that individual schools do not consistently favour either gender. Iceland also has the largest reading gender gap found in the PISA 2003 study. We hypothesise that as there appears to be a strong relationship between the size of the gender gaps in both reading and mathematics in all countries, the differences in mathematics could to some extent be dependent on reading ability. This needs to be explored further.

Nordic abstract

I dette kapitel ser vi på forskellene i de islandske PISA-resultater for piger/drenge og for by/land. Vi viser, at de usædvanlige islandske resultater, hvor piger klarer sig bedre i matematik end drenge, er et komplekst fænomen, som ikke umiddelbart lader sig forklare ved hjælp af de gængse teorier. En sammenligning mellem PISA 2003-resultaterne og de islandske nationale prøver for de samme elever bekræfter validiteten i PISA-resultaterne, men viser samtidigt, at kønsforskellene varierer betydeligt fra år til år og fra egn til egn, og således er vanskelige at forklare. Analysen viser også, at der ikke er noget fast mønster i, fra hvilke skoler henholdsvis piger og drenge klarer sig godt. En stærk sammenhæng mellem læse- og matematikfærdighederne kan også have haft en indflydelse på de islandske resultater.

In this chapter, the performance of Icelandic students in mathematics and reading in PISA 2003 will be examined. The variation in performance between urban and rural regions will be explored and compared to indications from the Icelandic National Examinations for the 10th grade, which show that girls perform better than boys in mathematics. In Iceland, girls generally also do better than boys in the PISA mathematics, reading, science and problem solving assessments. This is a unusual situation as the PISA results for other countries show that boys are generally better at mathematics, and therefore it is important to explore this further.

The literature on educational achievement abounds with results on gender differences in reading and mathematics (e.g. Caplan et al., 1997) and all kinds of social and cognitive explanations for the differences have been put forward. However, it is almost impossible to find any studies where girls do consistently *better* than boys in mathematics, although many researchers have shown that the well-documented gender gap in maths favouring boys appears to be diminishing (e.g. Walkerdine, 1998). A recent summary of international comparative studies indicates that the gender differences in mathematics achievement are generally rather small and in favour of boys (Stephens, 2003) and there are indications that a number of variables in the tests themselves - item format, reading load and text length, to name but a few - influence the results. However, none of the variables identified in the literature are really useful in explaining the reverse gender difference observed in Iceland. It is therefore important to understand this Icelandic phenomenon better.

Why then are girls so much better at mathematics in Iceland than in other countries? A possible interpretation for both the gender and urban-rural differences is the so-called 'Jokkmokk effect'. The Jokkmokk effect is a popular concept suggesting that boys in rural areas have values that prevent them from focusing on academic studies, while the girls in rural areas are perceived to see little hope for the future unless they concentrate on academic achievement, which ultimately enables them to move away and have a future elsewhere (see e.g. Ripley, 2005 for one example of this popular discussion). According to this explanation there should be a close link between urban-rural differences in academic performance and gender differences in performance.

Before attempting to verify the above phenomenon, we evaluated the reliability of the findings by comparing the results with other available data sets for the same population. The outcome of the initial analyses of the PISA data in Iceland led to an exploration of regional and gender differences in the Icelandic National Examinations for the 10th grade for the year 2003 and previous years. This was

done to test whether the gender differences in individual regions were consistent. The gender difference in a particular region must be observed consistently for several years at least, before searching for explanations for the difference in the particularities of the region. Potential explanations could be gender-specific unemployment, particularly attractive job opportunities for young men with few educational qualifications etc. In an attempt to explain the gender differences in Iceland, differences between schools were also explored, i.e. whether some schools consistently favoured girls or boys.

Delimiting regions

To compare the performance in mathematics and reading in urban vs. rural areas in Iceland, it was decided to divide the country into three regions which reflect the uneven geographical distribution of the population. Over half the population of Iceland lives in Reykjavik and its immediate surrounding area. This area is called the Greater Reykjavik Area, but is generally analysed as two separate areas in social research, namely Reykjavik and the Outer Reykjavik area. It could be argued, however, that these two areas are homogeneous with respect to life style and living conditions. Both units are urban and economically similar. Outside the Greater Reykjavik Area is the much larger sparsely populated area. It consists of small coastal towns and villages, scattered along the coast-line where people's livelihood has traditionally been based on fishing, and rural areas where farming is the main livelihood, and there are a few larger towns (from 4000 to 16 000 inhabitants) which form service and industrial centres for the more sparsely populated surrounding fishing and agricultural areas.

Thus, the division of Iceland into three regions: Reykjavik, the Outer Reykjavik area and the rural area is somewhat arbitrary and does not reflect any uniformity in the way of life or culture within each region. These geographical regions are therefore likely to need revision and more refined analyses that look beyond crude geographical divisions may be needed later. The population size of these three geographical regions is broadly similar. The number of 15-year-old students in Reykjavik in 2003 was 1090, while the Outer Reykjavik area contained 824 students and the rural area contained 1434 students.

Results: Performance in mathematics

Iceland is in 12th position compared to the other 41 participating countries, and in 9th position compared to the OECD countries. It should be noted that this ranking of countries is tentative, as the confidence intervals are quite large. Based on confidence intervals, Iceland is in 13th to 18th position among the 41

Table 1 *The performance of Icelandic students in Mathematics in PISA 2003 by geographical area*

		Math	SE	Space & shape	SE	Change and relation-ships	SE	Uncer-tainty	SE	Quan-tity	SE
Reykjavík	Girls	526	3.8	512	3.9	519	3.7	536	4.1	532	3.8
	Boys	518	4.1	505	4.2	512	4.3	538	3.8	509	4.3
	Total	522	2.8	509	2.8	515	2.7	537	2.5	520	2.7
Outer Reykjavik area	Girls	524	4.7	513	4.9	514	5.1	535	5.0	531	5.7
	Boys	516	4.1	505	4.6	515	4.9	535	4.8	508	4.8
	Total	520	2.8	509	3.0	514	3.0	535	2.8	519	3.1
Rural	Girls	520	3.4	509	3.8	511	3.8	527	3.8	523	3.6
	Boys	496	3.5	485	3.5	494	3.7	508	3.5	488	3.7
	Total	507	2.6	497	2.6	502	2.6	517	2.6	505	2.8
Total	Girls	523	2.2	511	2.3	514	2.3	532	2.4	528	2.3
	Boys	508	2.3	496	2.4	505	2.4	524	2.5	500	2.5
	Total	515	1.4	504	1.5	510	1.4	528	1.5	513	1.5

participating countries and in 8th to 16th position in comparison to the OECD countries. Among the Nordic countries, Iceland's ranking is average, not statistically different from Sweden and Denmark but better than Norway and worse than Finland.

Table 1 shows the mean performance in the PISA 2003 mathematics test overall and in each of the tasks for the three regions described above and the country as a whole (Björnsson et al., 2004).

The mean performance of 15-year-old students in Reykjavik in mathematics in PISA 2003 is 522 points, with a standard error of 2.8, which means that with 95% certainty the average performance is between 516 and 527. Within the mathematics tasks, the students' performance is more varied. As the table shows, the performance is best on the *uncertainty* tasks (537 points), based on questions related to statistics and probability. These are also the tasks where the gender difference is smallest.

The table also shows that there is very little difference between Reykjavik and the Outer Reykjavik area in mathematics. The total scores are 522 and 520 respectively. This is not surprising, given the similarity in overall lifestyle in these

two adjacent geographical regions. However, the performance in the rural area is markedly lower, with a total score of 507. It is of particular interest to note that girls in rural areas perform no worse than girls in the two urban areas and even have higher average scores in some cases. It is the boys who bring down the average in the rural area, scoring only 496 compared to 520 scored by girls in the rural area and 516 and 518 scored by boys in the two urban areas.

Thus it turns out that in PISA 2003 the gender difference in performance is much larger in the rural area, and is not significant in the two urban areas. This requires an explanation.

Unusual gender difference in Iceland

The usual outcome in other countries is that girls always perform better in reading than boys, and boys generally perform better in mathematics than girls. This pattern is not observed in Iceland. The gender difference in reading is larger in Iceland than in any other country in PISA 2003 and Iceland is the only country where girls have a considerable advantage over boys in mathematics. The girls have an overall 15 point advantage over boys. The country most similar to Iceland in this respect is Thailand, where girls have a 4 point advantage over boys, which is not significant. The four countries with the smallest gender difference in favour of boys (albeit non-significant) are not OECD countries: Serbia, Latvia, Indonesia and Hong Kong-China.

In the other Nordic countries the tendency is for boys to perform better in mathematics. This tendency is strongest in Denmark (17 points), but in Sweden, Finland and Norway it is low (7, 7 and 6 points respectively; OECD, 2004).

Table 2 shows the gender difference in the mathematics test scores overall and for the individual tasks in the captial Reykjavik, the Outer Reykjavik area and the rural area. The table shows the points advantage for either gender for each of the mathematics tasks.

Significant gender differences are only observed in rural Iceland, with the exception of the *quantity* tasks, where the difference is marked and significant everywhere in the country. Boys are better than girls in only two cases: in Reykjavik they perform better in the *uncertainty* tasks and in the Outer Reykjavik area they perform better in the *change and relationships* tasks. However, these differences are small and non-significant.

Table 2 *Gender difference in mathematics by geographical area*

		Maths total	Space & shape	Change and relationships	Uncertainty	Quantity
Reykjavík	Girls	8	7	6		23*
	Boys				2	
Outer Reykjavik area	Girls	8	8		0	23*
	Boys			1		
Rural	Girls	25*	25*	17*	18*	35*
	Boys					
Total	Girls	15*	15*	9*	7	28*
	Boys					

*p<= 0.05

Proficiency levels

Urban-rural and gender differences in the PISA 2003 study can also be viewed in terms of the number of students who fall into each of the six proficiency levels. It can be assumed that individuals who fall into levels 0 and 1 are unable to use mathematical methods and apply mathematical understanding in daily life to any significant extent. In figure 1, we see the distribution of boys' and girls' proficiency levels across the three regions.

Figure 1 *Percentage of students at each proficiency level by region*
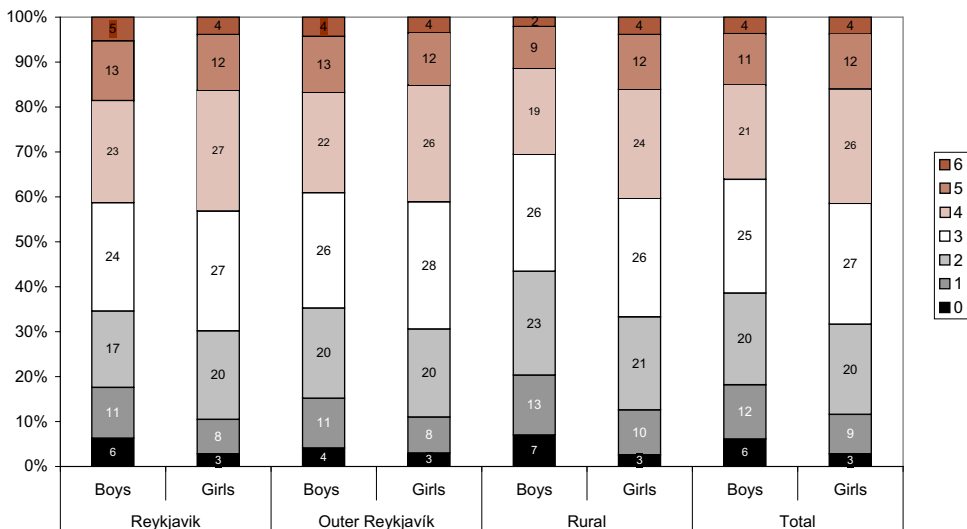
Figure 1 shows that there is a considerable difference in proficiency levels between regions and by gender. Interestingly, Reykjavik is similar to the national average with 6% of boys and 2.8% of girls falling into level 0. In the Outer Reykjavik area, the difference between genders is smallest. At the other end of the scale (level 6) far more boys are ranked at level 6 in Reykjavik than in the other two areas. Reykjavik thus seems to contain a large proportion of both low and high achievers. The percentage of poorly performing boys is, however, greatest in the rural area, with 7% of the boys ranked at level 0. The picture of regional differences is therefore similar, whether we look at proficiency levels or at test averages.

If the gender differences in proficiency levels by region are examined, it is clear that in both urban areas, there is a greater proportion of boys ranked at level 6 compared to girls. However, in the rural area, there are more girls in level 6 than boys.

Other PISA subjects

There is a similar gender difference in all PISA 2003 subjects in Iceland. Girls score 15 points higher than boys in mathematics. In reading, girls score 58 points higher than boys, which is the largest difference in any of the participating countries. In science, the girls' advantage is 10 points (equal to Tunisia, with no other country showing a higher difference) and in problem solving the difference is 30 points, by far the greatest difference favouring girls in the whole study.

The performance of Icelandic students in reading, broken down by region and gender, can be seen in table 3 below.

Table 3 *Performance in Reading and gender difference by region*

		Reading	SE	Gender difference - points -
Reykjavík	Girls	529	3.7	
	Boys	469	4.2	
	Total	498	2.6	60*
Outer Reykjavik	Girls	518	4.2	
	Boys	470	5.2	
	Total	493	3.3	48*
Rural	Girls	518	3.2	
	Boys	457	3.4	
	Total	486	2.5	61*
Total	Girls	522	2.2	
	Boys	464	2.3	
	Total	492	1.6	58*

*p<= 0.05

Table 3 shows that there is a very large difference between the genders in favour of girls in all three regions. There is no difference between rural and urban areas. Whatever factors influence performance in rural and/or urban areas seem to be affecting the subjects in different ways.

Are these results reliable?

A number of explanations are possible for the gender and regional differences observed and many lines of research can be considered depending on theoretical inclinations. How can we explain the strong and unusual gender difference favouring girls recorded in Iceland in all subject areas tested in PISA 2003? We mentioned earlier the 'Jokkmokk effect', according to which the boys in rural areas are attracted by well paid jobs and away from learning, while the lack of similar opportunities for girls steers them towards pursuing their academic studies. In the light of the findings above this theory deserves closer attention.

But first of all, let us assess the reliability of the above findings. We will focus on mathematics, since it is the main subject area in PISA 2003. Are girls in rural areas systematically outperforming boys year after year? To assess the stability of the regional gender difference over several years we analysed the Icelandic National Examinations (INE) for 15-year-olds as this allows yearly comparisons, whereas there is a 3 year interval between PISA studies.

For this analysis the concordance between the PISA results and the Icelandic National Examinations for 10th grade students was explored, in order to discover whether we were justified in using the Icelandic National Examinations to test the consistency of the gender difference found in PISA. The aim of these two tests is not identical and the rationale behind them is different. However, one would expect that mathematics tests in both PISA and the Icelandic National Examination assess to a large extent the same thing. In figure 2 the results for both tests are compared by geographical region. For the purposes of this comparison, the rural region is broken down into smaller areas for a more detailed analysis. Figure 2 shows the performance of girls and for boys in the two tests by region.

We see that the performance of students overall and the gender difference are similar in each region for both tests in 2003. This figure shows that there is a fair agreement between the PISA and Icelandic National Examinations assessments for 15-year-olds, giving mutual concurrent validity to both tests.

Having established that there is a fairly good agreement between the PISA and Icelandic National Examination results in the ranking of performance by region, we can move on to explore the reliability of the gender differences over several years

Figure 2 *Comparison between gender difference by region on PISA 2003 and the Icelandic National Examination 2004*
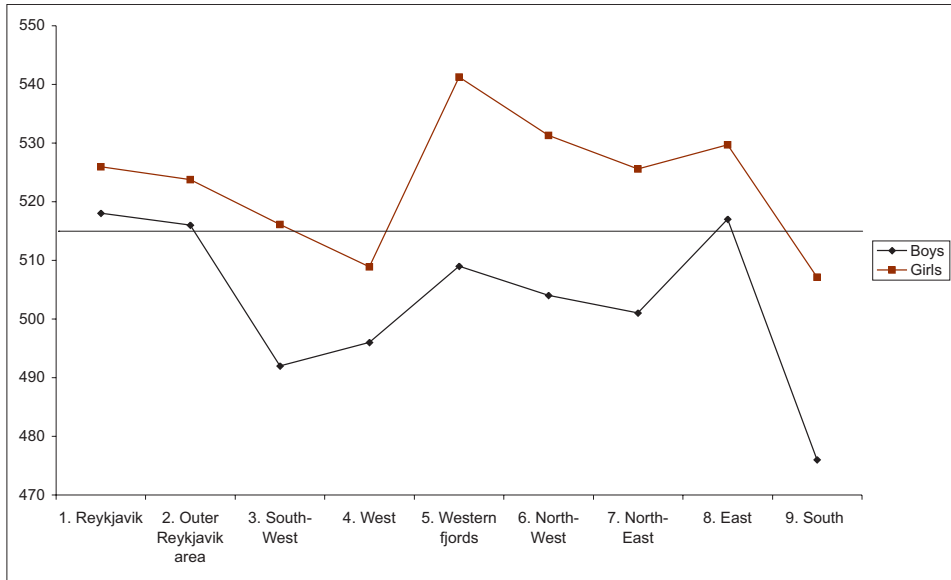
Mean performance of girls and boys on the PISA 2004 Mathematics scale by region
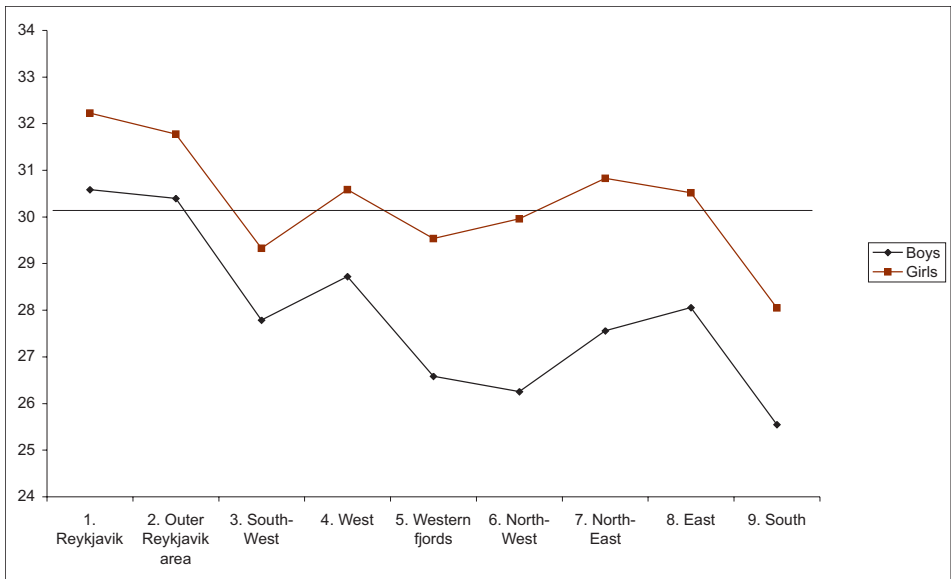
Mean performance of girls and boys on the 2004 Icelandic National Mathematics Test by region

in different regions. Figure 3 shows gender difference by region in the Icelandic National Examinations from 1996 to 2004.

Figure 3 *Gender difference on the Icelandic National Examinations 1996-2004 by region*
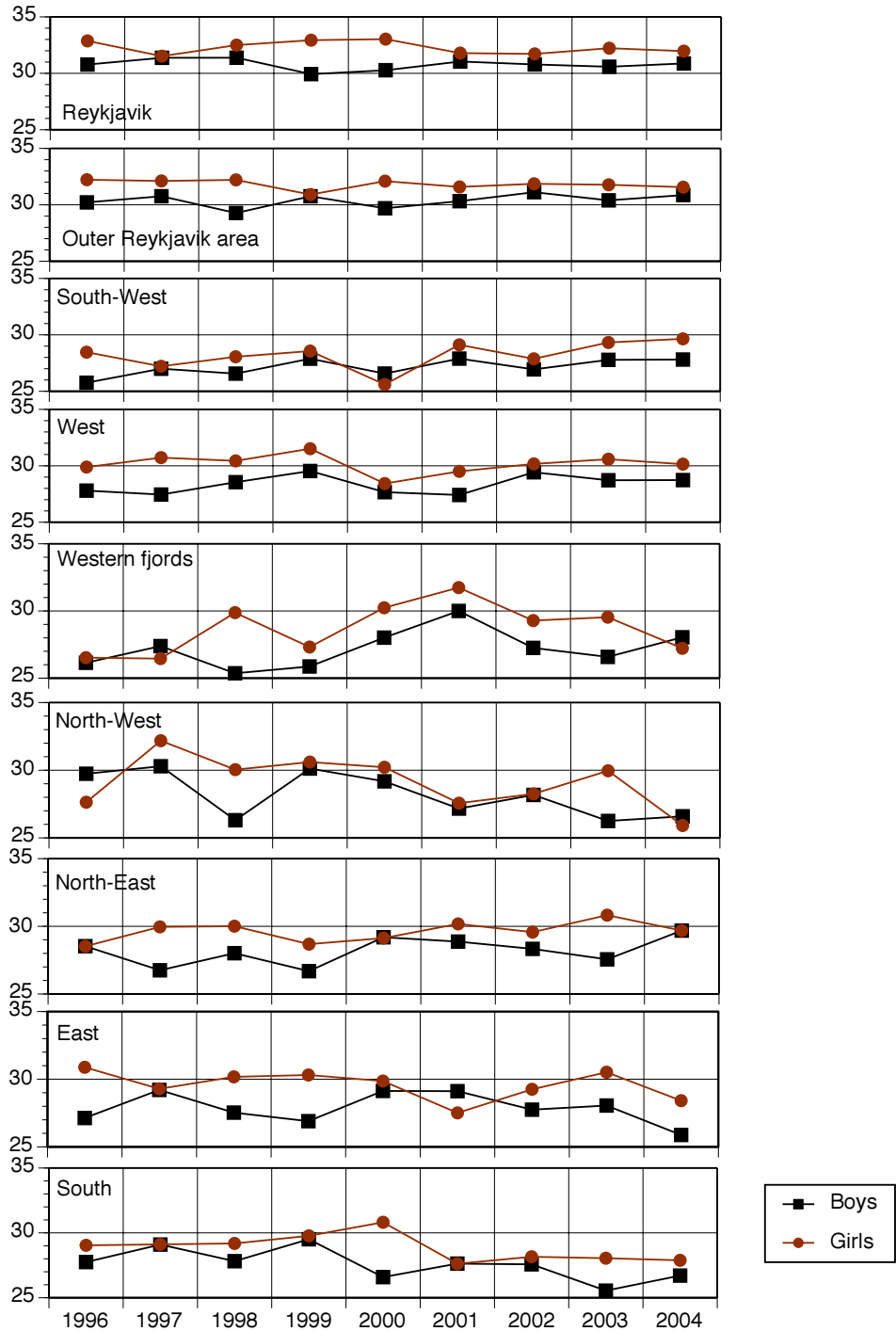
Figure 3 shows that there is very little similarity over the years between the performance of girls and boys in each region. Thus, the findings linking gender differences to particular regional characteristics are not consistent. Even when rural areas are compared to urban areas gender differences are not consistent over time. The years 2000 and 2003 are a good example of this inconsistency. In 2000, there is very little gender difference in the rural areas but a marked difference in the urban areas, while in the year 2003, the situation is reversed.

The school effect?

To explore whether some schools could be said to favour one gender over the other separate analyses were conducted on the Icelandic National Examination results for 10th grade students from 1996 to 2003. The gender difference in each school was computed, with the male average in each school being subtracted from the female average in the same school. This produced eight variables, one for every year, showing the degree to which each school favoured girls that particular year. These eight variables were then correlated using Spearman's correlation. The assumption was that if something about a particular school favoured either gender one year, it would be likely that this would be repeated the following year and in later years, thus yielding a positive correlation between years.

However, as table 4 shows, there is only a significant correlation between 2 pairs of years and one of these correlations is negative. Such correlations are likely to be

Table 4 *Correlation of gender difference between schools in 1996 – 2003*

	1997	1998	1999	2000	2001	2002	2003
1996	-0.002	-0.126	-0.79	-0.166	-0.06	-0.052	-0.061
1997		0.042	-0.118	-0.029	0.193*	-0.112	0.050
1998			0.077	0.088	0.145	0.179	-0.007
1999				-0.039	-0.052	0.009	0.118
2000					0.021	-0.049	-0.026
2001						0.058	-0.055
2002							-0.207*

* Correlation is significant at the 0.05 level (2-tailed).

obtained by chance. Overall, the results indicate that there is no consistency in the degree to which a school favours either gender.

The relationship between reading and mathematics

When viewing the overall PISA 2003 and PISA 2000 results it is clear (OECD, 2003) that a very strong correlation between performance in reading and mathematics exists. The correlation in PISA 2003 is around 0.6 for most countries and slightly higher for Iceland. This means that students who are better at reading are also generally better in mathematics. Results show that when the maths performance of boys and girls is compared, if the effect of reading ability is controlled for, then, unexpectedly, the Icelandic boys are a little better at mathematics than the girls. In other words, given equal proficiency in reading, boys can be expected to be slightly better at mathematics than girls. This relationship between reading and mathematics is therefore an important factor, suggesting that looking at the maths gender differences in isolation from performance in other subjects is perhaps pointless. As Iceland has the largest gender gap in reading in PISA 2003, the above correlation would therefore predict that girls should perform strongly in mathematics compared to boys. This appears to hold true for most of the participating countries in PISA generally; when the gap favouring girls in reading is smaller the boys generally perform better in mathematics. In Iceland where the gender gap in reading favouring the girls is large, we should therefore also see the smallest gender difference in maths favouring boys or indeed a gender difference in favour of girls, as was the case in PISA 2003. This simple relationship between reading and mathematics is of course not isolated from or uninfluenced by all the background variables that have been implicated in gender differences. The relationship is also not completely linear but it appears clear that it is very strong and perhaps stronger than any other relationship used to explain the gender differences in educational achievement. For further discussion of this issue, see the chapter by Roe and Taube (this volume). This relationship can therefore also be used to explore the regional differences examined earlier, as the reading gap is generally also bigger in rural than in urban areas. But this neads further research and is outside the scope of this chapter.

Gender differences at the international level

Our attempts to analyse gender differences favouring girls in the PISA 2003 mathematics test results in Iceland at the national level or in terms of regional or school differences have not yielded conclusive results. However, further avenues for research can be pursued at the international level. The results of a number of international studies can be correlated with the gender differences observed in PISA

across countries. Our initial analyses of these studies have found that gender differences in favour of girls are associated with positive measures of women's empowerment: a high level of women's empowerment correlates with a better performance by girls in comparison to boys within each country. Similarly, gender differences can be correlated with international measures of corruption, democratic development, economic situation etc. There is much research to be done in these areas in order to understand the unusual Icelandic gender gap, and the results of this research will be reported in a separate publication.

Conclusion and discussion

Overall, the results of these analyses can be summarised as follows. We find that the regional difference and gender difference in PISA 2003 results are replicated within the same year in the Icelandic National Examinations for the 10th grade. This finding *per se* is evidence for the validity of both tests. In addition, this similarity justifies our use of National Examination results for other year groups in our search for an explanation of the PISA results.

However, we find that gender differences in each region in the Icelandic National Examination are inconsistent from one year to the next, which makes it difficult to associate gender differences with particular regional characteristics. Also, we find that there is no consistency within schools as to whether they favour boys or girls. A school which shows girls to be better one year may show the opposite the next year.

Our attempts to explain why boys perform worse in rural Iceland than girls and than boys in the rest of the country have thus far not proved fruitful. There is great variability in gender differences within rural and urban areas and between years and schools. Unless gender and regional differences are found to be consistent across subjects and over time, there is little point in looking for established regional characteristics to explain the differences observed in one year. With the addition of the PISA 2006 results for *Link items* which will be published in 2007, results from 2000, 2003 and 2006 for mathematics, reading and science can be combined on a 3-point timeline and trends in gender differences using the PISA results can be studied more reliably and compared to the results presented in this chapter for the Icelandic National Examinations.

We believe that a combination of approaches, local and international, is necessary to throw further light on these gender and regional differences. The inconsistency in the differences in performance analysed by gender and region does indicate that gender differences in Iceland cannot be explained by the simple 'Jokkmokk effect'

and that there are considerably more complex relationships at work. Over the past few decades, migration between certain parts of the country has increased considerably, mostly characterised by movement from rural to urban areas. This has surely had an effect but it is very difficult to relate this effect directly to the gender differences observed in the PISA results. During the last 10 years there have also been considerable socio-economic changes in Iceland, with the population generally doing better economically. These changes have not yet been related to outcomes of educational measurements, but we would expect to see a correlation between these variables.

References

Björnsson, J.K., Halldórsson, A.M. & Ólafsson, R.F. (2004). Stærdfrædi vid lok grunnskóla: Stutt samantekt helstu niðurstaðna úr PISA 2003 rannsókninni. Reykjavik: Námsmatsstofnun. (Icelandic PISA 2003 Report, Educational Testing Institute).

Caplan, P.J., Crawford, M., Hyde, J.S & Richardson, J.T.E. (Eds). (1997). Gender Differences in Human Cognition. New York: Oxford University Press.

OECD. Learning for Tomorrows World. First results from PISA 2003. OECD Paris.

Ripley, A. (2005). Who says women can't be Einstein. Time magazine March 7, 2005.

Roe, A. & Taube, K. (chapter 10 this volume). How can reading abilites explain differences in maths performances?

Stephens, M. Gender Differences in Achievement. In J.H. Moskowitz & M. Stephens. (2003). *Comparing Learning Outcomes*. International Assessment and Educational Policy. London and New York: RoutledgeFalmer.

Walkerdine, V. (1998). Counting Girls out: Girls and Mathematics. London: Falmer Press.

Chapter 15

Leaving Examination Marks and PISA results – Exploring the Validity of PISA Scores

Jan Mejding, Simon Reusch and Thomas Yung Andersen

Abstract

Through an established link between PISA scores and leaving examination marks for the Danish sample of pupils taking part in PISA 2003 it has been shown that there is a significant correlation between the two, even though the purpose and the structures of the tests are very different in the two test situations. This correlation strengthens the notion that PISA measures substantial competencies and does not just reflect a learned test taking competency.

Nordic abstract

Gennem en sammenkobling af karakterer fra folkeskolens afgangsprøve med PISA-resultater for de elever, som deltog i PISA 2003 undersøgelsen i Danmark vises det, at der er en signifikant sammenhæng mellem elevernes afgangsprøve resultater og de samme elevers resultater i PISA testen. Den fundne sammenhæng er tilfredsstillende ikke mindst set i lyset af at folkeskolens afgangsprøve og PISA testen har forskellige udgangspunkter og formål. Undersøgelsen styrker således opfattelsen af at PISA måler noget substantielt kompetence-/færdighedsmæssigt og ikke blot er et udtryk for elevernes tillærte testkompetencer.

Will a young person who did well in PISA also do well when confronted with the leaving examinations or is there no connection between the skills measured in PISA and the skills the pupils need in school? This question will be investigated in this article.

In this article we talk about PISA results in the plural because PISA consists of – at the least – three different subject areas: reading, mathematics and science. In the context of PISA these areas are defined not by the expectations of a given

curriculum but by the kind of skills expert groups judge to be important for young people entering further education and the labour market.

It is therefore of interest that through our access to the Danish examination and PISA data we have had an opportunity to establish a link between PISA 2003 results and the marks of the pupils taking the leaving examination of the Folkeskole (the municipal primary and lower secondary school).

The first question, however, is whether we should expect any connection between the PISA score and the leaving examination marks.

As stated above, the point of the development of PISA was quite a different from that of the leaving examination. The latter is aimed at measuring how well pupils have mastered specific subject curricula and the topics taught in school. PISA on the other hand aims to measure how good pupils are at using the knowledge and the skills they have learned in school in applied situations that simulate problem settings from real life. And these skills need not have been included in the school's curriculum. We might expect therefore that some pupils have been successful at learning the subjects taught in school but have problems applying what has been learned in a different situation. Their learning is context-dependant and is only useful within the school setting. At the same time it is true that not all the skills that PISA evaluates are included in national curricula across the world.

The construction of the PISA score

There is a big difference in the way 'points' are awarded in PISA and in the leaving examination. In PISA correct answers to questions in the different subject areas are counted. The questions are not all of the same difficulty and through an item analysis the answers are 'weighted' according to their difficulty level. All pupils do not answer the same questions; they only answer some of the questions that as a whole constitute the score for the total population. On the basis of all the pupils' answers and through different statistical procedures a national score is computed for each subject area. Thus PISA is constructed to give a precise population estimate and not an adequate measure for each single pupil[1].

---

1. In fact when computing a population's performance five different random numbers are drawn from the distribution of scores that could reasonably be assigned to each individual pupil and it is the average of these five 'Plausible Values' (PVs) that constitutes the basis for the calculation of a national mean. However, this analysis is concerned with the relation between individual pupils' performances in PISA and their marks in the leaving examination and that is quite a different kind of analysis. We therefore use the so-called Warm's Weighted Likelihood Estimator (WLE) score when conducting this analysis. These scores can: '.. be treated as (essentially) unbiased estimates of student abilities, and analysed using standard methods.' (Adams & Wu, 2002, p. 106).

Marking

The marks for the leaving examination have a totally different purpose (see a description of the Danish marking scale in table 1). Marks in a leaving examination should indicate how well the pupil has mastered the relevant curriculum. Even though the marks given should be criterion related, i.e. on the basis of a description of what should be mastered within a given subject, it is often a problem that subjective judgement plays an important part: Pupils regarded by the teacher as good at learning are more readily given higher marks and so forth. Teacher's concepts of marking very often also include a – more or less unconscious – reference to a norm: the normal distribution. Most marks then will fall in the middle of the scale and fewer at either end. But an individual class is not a valid representation of the total population and therefore this tendency to normalise marks may give an inaccurate picture of where these pupils actually are in relation to the rest of the population. A particular mark, e.g. 8, which is the middle mark on the Danish marking scale, will not be given for the same performance by every teacher, and it has been shown that there are rather large differences between the marks that different teachers give for the same performance (see Miller, 2004; Linde, 2003 and Gregersen, 1984 for a broader discussion on this issue). In the Folkeskole leaving examination marks are given jointly by the pupils' teacher and an external examiner. This will to some extend minimize the influence of subjective factors on the marking, but even then it is questionable how precisely a given mark represents a

Table 1: *Marking scale in the Danish Education System*

13: Is given for the exceptionally independent and excellent performance.

11: Is given for the independent and excellent performance.

10: Is given for the excellent but not particularly independent performance.

9: Is given for the good performance, a little above average.

8: Is given for the average performance.

7: Is given for the mediocre performance, slightly below average.

6: Is given for the just acceptable performance.

5: Is given for the hesitant and not satisfactory performance.

03: Is given for the very hesitant, very insufficient and unsatisfactory performance.

00: Is given for the completely unacceptable performance.

pupil's competency in a subject. In some subjects, i.e. spelling and written mathematics, it is easier to give an objective description of what should be learnt, while in other subjects, such as in oral Danish, the subjective impressions of the marker will play a stronger role.

In spite of these differences the Danish national reports for both PISA 2000 and PISA 2003 argue that a large proportion of the competencies that PISA aims at measuring are in fact a part of the national curriculum in the tested subject areas. This is especially true for mathematics and science, while reading constitutes a meta-competency that plays a role as much in the study of Danish as in mathematics and science, since reading is also one of the skills needed to solve the questions within these subject areas. But even for reading on its own it is true that many of the reading skills targeted in PISA are also described as important competencies in the national curriculum for Danish.

Hypothesis

We would generally expect to find a correlation between a pupil's PISA score and his or her leaving examination mark. But we would also expect there to be a dispersion of PISA scores within each marking category as the two 'scores' have different purposes and therefore will measure the same competences only to a certain extent.

Method

The starting point for our analysis is the Danish PISA 2003 student sample for 15^1/$_2$-year-olds. In this sample there are 4215 respondents. The PISA scores were linked through a unique student-id to the UNI-C LEM-database (of leaving examination marks) for the different subjects. It should be noted that the leaving examination marks do not necessarily originate from the same year (2003) as the PISA results, as some pupils started school earlier or later than the normal age for starting school - which in Denmark is the year in which a child turns seven. Some of the pupils even finished school with the advanced leaving examination (taken at the end of grade 10) instead of or as a supplement to the leaving examination (taken at the end of grade 9). In 2004 it was possible to identify the leaving marks of 3729 pupils who also took part in PISA. Of these 3729 pupils 1% took the advanced leaving examination (grade 10) and 1.9% took their examination in 2002 (one year earlier than normal), 86.6% took their examination in 2003 (at the normal time) and 11.5% in 2004 (one year later than normal).

Table 2 shows that the mean marks for this sample are almost identical to the national mean in 2003, when the majority of PISA students took the leaving examination.

Table 2 *Overview of Leaving Examination Marks for the Danish PISA 2003 Sample*

	Mean	Std. Deviation	Number	National mean in 2003
LEM Danish - spelling	7.76	1.527	3674	7.7
LEM Danish - written	8.00	1.372	3672	8.0
LEM Danish - oral	8.52	1.667	3671	8.5
LEM Mathematics - written	7.87	1.536	3671	7.9
LEM Mathematics - oral	8.25	1.553	3660	8.2
LEM Physics-Chemistry	7.90	1.664	3324	8.0

Results

Distribution of PISA scores within LEMs

When the leaving examination marks are compared with the related PISA score there is generally a satisfactory agreement between the two. Pupils who gained higher marks in the leaving examinations generally had higher PISA scores as well.

Figure 1 *Marks in written Mathematics and PISA Maths scores*
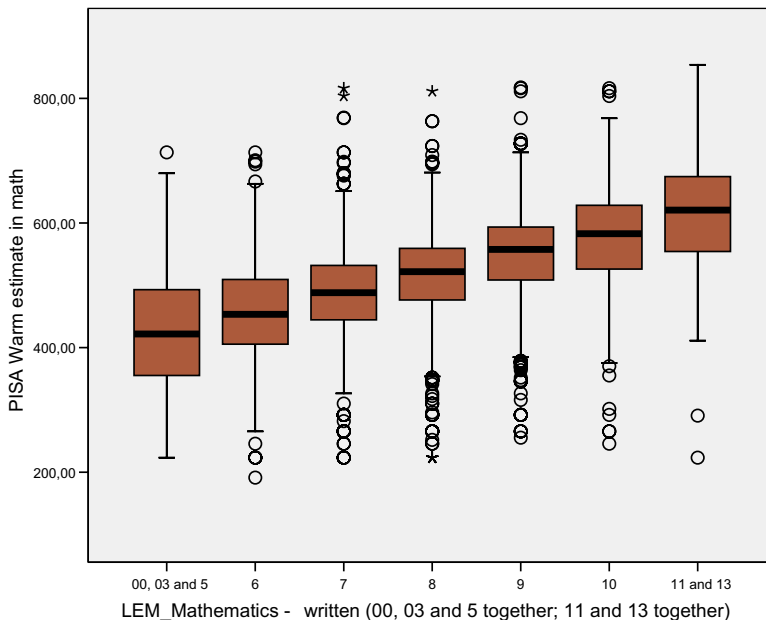
Figure 2 *Marks in oral Mathematics and PISA Maths scores*
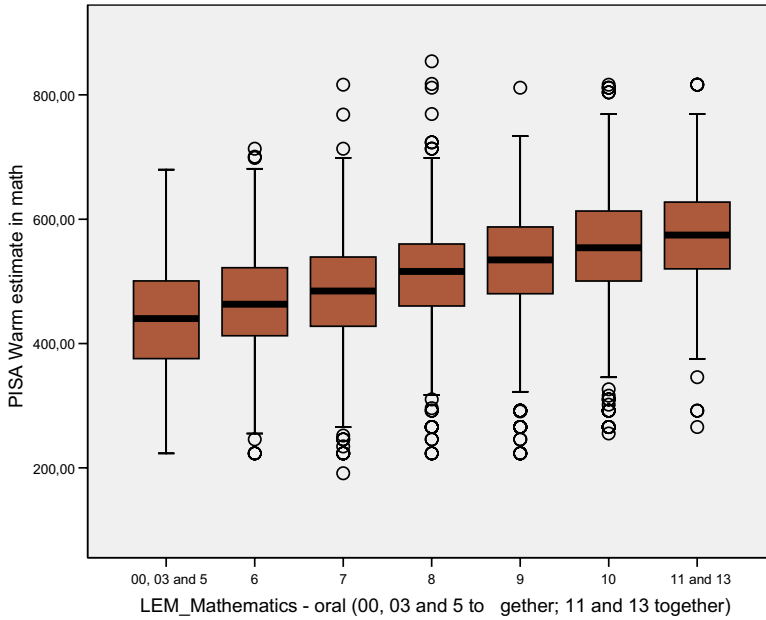
Figure 3 *Marks in Danish Spelling and PISA Reading scores*
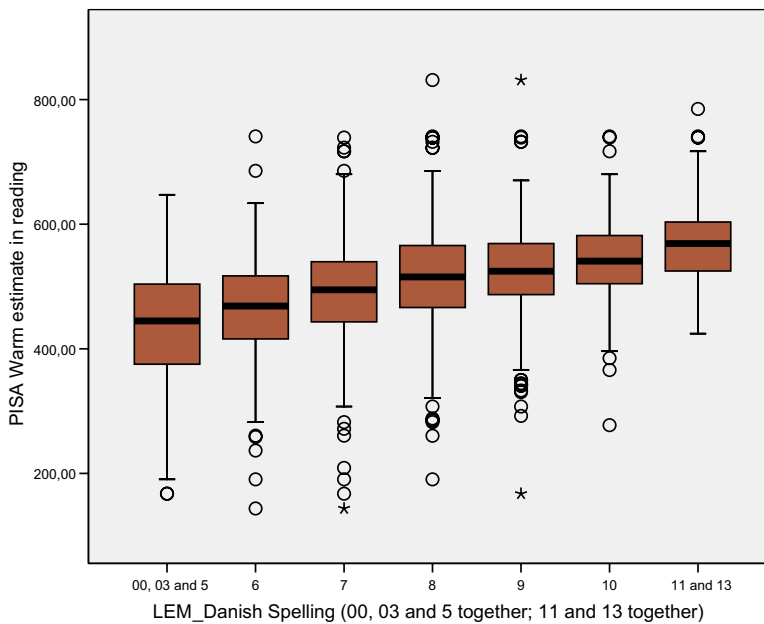
Figure 4 *Marks in oral Danish and PISA Reading scores*
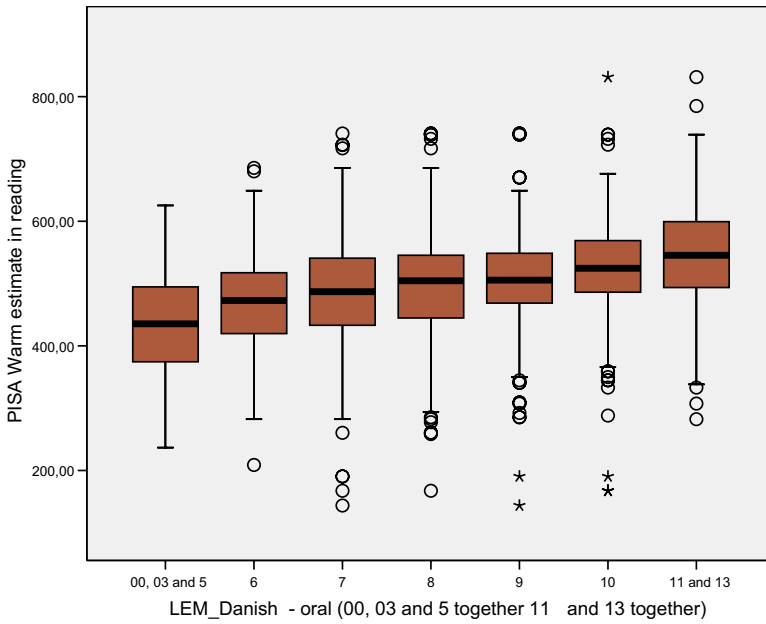
Figure 5 *Marks in Physics and Chemistry and PISA Science scores*
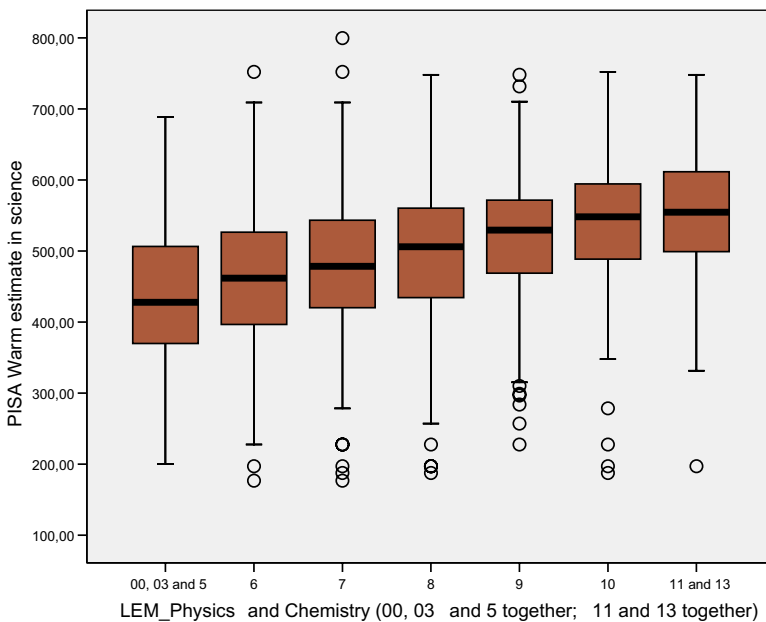
Figure 6 *Marks in oral Mathematics and PISA Science scores*
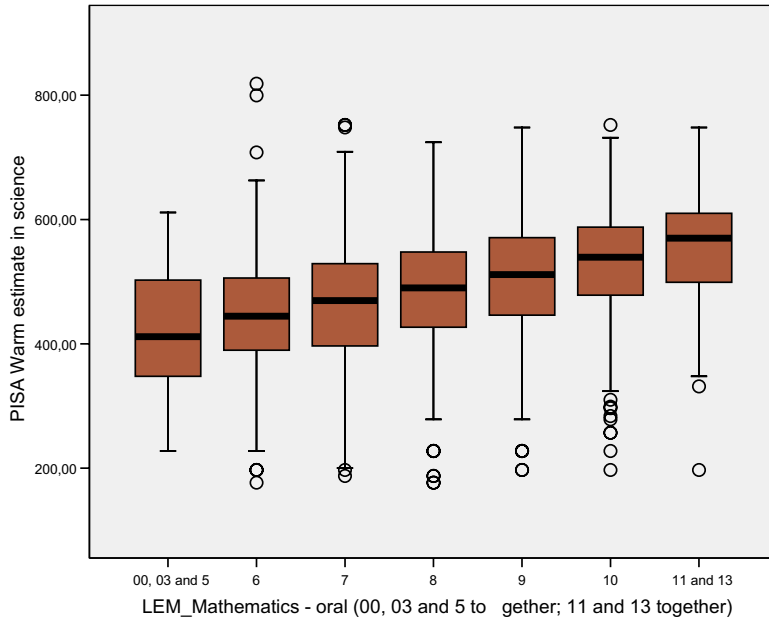
The figures[2] 1 to 6 show this agreement between the different PISA scores and the examination marks from related subject areas. The marks at the extremes (0 to 5 and 11 to 13) have been collapsed as the number of pupils in these categories was very small.

The relationship between the PISA scores and the leaving examination marks can be shown through bivariate correlation coefficients (Pearson's R). Table 3 shows correlations between PISA domains and the leaving examination marks; table 4 shows correlations between different leaving examination marks and table 5 shows correlations between the different PISA domains.

Correlations between leaving examination marks

It can be seen from table 4 that the correlations are strongest between related subject areas: The correlation between oral and written mathematics and between spelling and written Danish is 0.68.

---

2. The figures are so-called box plots. The box shows the distribution of the middle 50% (25%-75%) around the median (the black line). The thin lines and the small circles show the rest of the distribution (0%-25% and 75%-100%).

Table 3 *Correlations between Leaving Examination Marks and PISA results*

Correlations	PISA-maths	PISA-probl	PISA-read	PISA-science
Danish - spelling	0.34	0.30	0.37	0.32
Danish - written	0.28	0.24	0.30	0.29
Danish - oral	0.27	0.23	0.28	0.27
Mathematics - written	0.46	0.38	0.34	0.43
Mathematics - oral	0.36	0.29	0.27	0.32
Physics-Chemistry	0.31	0.25	0.28	0.30

All correlations are significant at the 0.01 level (2-tailed)

Table 4 *Correlations between Leaving Examination Marks*

Correlations	Danish - spelling	Danish - written	Danish - oral	Maths - written	Maths - oral
Danish - written	0.681				
Danish - oral	0.47	0.53			
Mathematics - written	0.56	0.47	0.44		
Mathematics - oral	0.46	0.44	0.46	0.685	
Physics-Chemistry	0.38	0.38	0.42	0.57	0.57

All correlations are significant at the 0.01 level (2-tailed)

Table 5 *Correlations between PISA Domain results*

Correlations	PISA-maths	PISA-probl	PISA-read
PISA-probl	0.66		
PISA-read	0.53	0.55	
PISA-science	0.62	0.58	0.62

All correlations are significant at the 0.01 level (2-tailed)

The third highest correlation is somewhat lower: The correlation between physics-chemistry and mathematics (both oral and written) is 0.57 – but these subject areas can be regarded as being related as well. There are correlations between spelling and written mathematics of 0.56 and between the related areas written and oral Danish of 0.53. All other correlations here are below 0.50.

Correlations between PISA domains

The strongest correlation between PISA domains is 0.66 which is found between mathematics and problem solving. The lowest correlation (0.53; see table 5) is

found between mathematics and reading. These relatively strong correlations between the four different PISA domains can probably be explained by two circumstances.

First of all the different PISA scores are derived from test units that are presented as one test, and the construction of the test units themselves share similarities in the way they are constructed and presented to the pupils. This eliminates influences from fluctuations in pupils' day-to-day performance related to energy levels and illness or from disturbances from external sources such as high or low temperatures, noise etc. Together with the more uniform test format between domain areas, this type of presentation will tend to heighten correlations. The Leaving examinations, on the other hand, are held on different days and involve quite different test formats and are therefore more likely to be influenced by external factors not related directly to the competencies in question.

Second, the PISA test is a written test and scores are meticulously marked using a uniform procedure across countries. The oral leaving examinations take place in quite a different setting where the dialogue between pupil and teacher is essential but where too the teacher's and the external examiner's subjective impressions of the pupil, as well as their personal beliefs and values, play an important part in the evaluation of the pupil's performance.

Correlations between PISA domains and leaving examination marks

The relation between the PISA domains and the leaving examination marks can be seen from table 3 and by following the correlation curves for each PISA domain shown in figure 7. These curves show that the leaving examination mark for written mathematics is strongly correlated not only to PISA mathematics but also to the other three PISA domains. This means that a pupil who masters the competencies measured by the leaving examination in written mathematics to a large extend also will master the competencies needed in all PISA domains. However, the same relationship is also found between the leaving examination marks in mathematics, physics-chemistry and Danish spelling. It seems that if a pupil does well in written mathematics then this pupil will do well in other subject areas too.

Both in Danish and in mathematics the PISA scores are more strongly correlated with the written leaving examination marks than with the oral marks. This is to be expected, as both test formats are written. Thus the competencies measured by the oral leaving examinations are generally less credited in the written PISA test.

The overall reading literacy score (PISA-read) is correlated most strongly to the subject areas of Danish and written mathematics, but it is also correlated with the other leaving examination marks, as reading is a general competency needed in all subject areas.

Figure 7 *Intercorrelations*
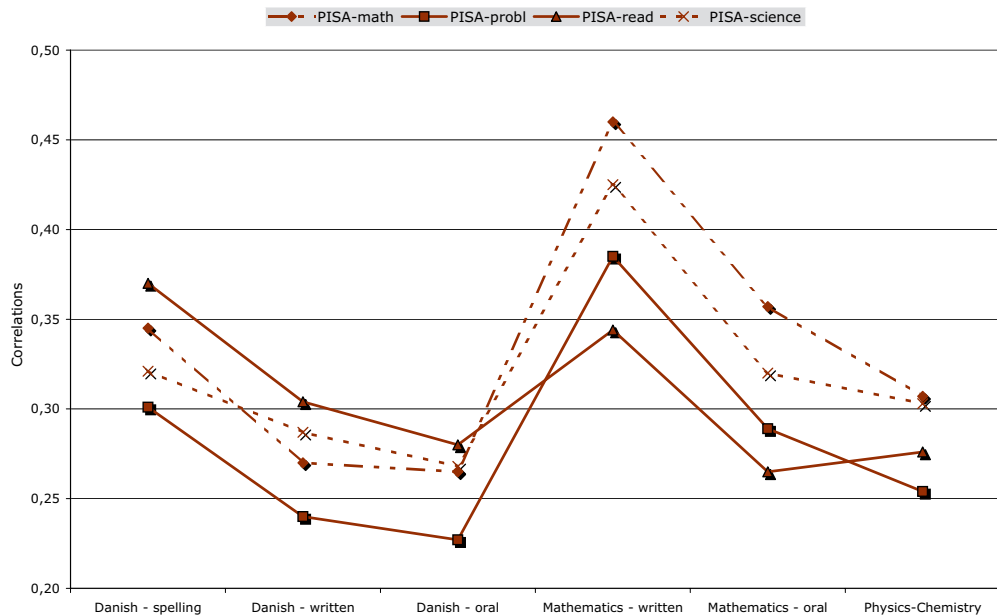
It is noteworthy that the correlations between scores for problem solving in PISA and the leaving examination marks are rather low, in spite of the relative high correlations between all four PISA domains. This seems to indicate that the PISA problem solving tasks measure competencies other than those generally reflected in the leaving examination marks.

Discussion

The leaving examinations and the PISA domain scores have been constructed for different purposes. In spite of this we find a relatively strong correlation between mean PISA scores and each marking category within the different subject areas of the leaving examination. This correlation strengthens the notion that PISA measures substantial competencies and not just the competency of learning how to take a test.

It has been shown previously (M.N. Hansen & A. Mastekaasa, unpublished data) that there is a significant relation between the marks that pupils themselves report that they have been given lately - and PISA scores. Correspondingly, Dines Andersen showed (Andersen, 2005) that there is a significant relation between a pupil's PISA

score in reading and the same pupil's chances of having completed his/her youth (upper secondary) education. The lower the PISA reading score, the higher the risk of dropping out (Andersen, 2005 p. 9).

Two other Nordic studies have looked at the relation between PISA scores and national test results: In Sweden in 2000 participating schools were asked to list the pupils' marks in Swedish. These marks were given at the end of the school year by the teachers. The marks correlated significantly with the PISA 2000 reading score, but there was also a considerable spread in the distribution. (Skolvärket, 2001, pp..51-52). In Iceland in 2003 the PISA reading scores were compared with results from the Icelandic Language National Standard Test. This test was given to all the pupils at the end of the school year – approximately one month after the PISA test – and like the PISA test this test is a written test targeting a number of competencies within language skills. The correlation between these two tests was as high as 0.6 – a highly significant result.

Conclusion

In the present study we have shown that in Denmark also there is a strong relation between a pupil's PISA score and the same pupils leaving examination marks. However, as would be expected, there is significant overlap between score levels and marking categories. Any given PISA score is therefore not unambiguously translatable into a marking category and vice versa. Finally, PISA has been constructed in order to monitor competencies in relation to the future demands of society. An analysis of the educational progress of the present sample will show if it is also possible to demonstrate a significant correlation between a given pupil's PISA score and his/her future educational course for the PISA 2003 generation.

Literature

Adams, Ray and Margaret Wu (eds*): "PISA 2000 Technical Report"*, OECD, 2002

Andersen, A.M. et al: *"Forventninger og færdigheder – danske unge i en international sammenligning"*, SFI-SURVEY, København, 2001

Andersen, Dines: *"4 år efter grundskolen – 19-årige om valg og veje I ungdomsuddannelserne"*, AKFforlaget, København, 2005

Bjørnsson, JK, Halldorsson, A, & Olafsson RF. *"Stærdfræði vid lok grunskóla: Helstu nidurstödur úr PISA 2003 rannsókninni".* Námsmatsstofnun 2004. (PISA 2003 Icelandic National Report, Publisher: Educational Testing Institute 2004).

Gregersen, Jørgen: *"Karakterer & kommunikation"*, DLH, København, 1984

Hansen, Marianne Nordli and Arne Mastekaasa: *"The Nordic Welfare States and Equality of Educational Opportunity"* (unpublished: Department of Sociology and Human Geography, University of Oslo)

Linde, Göran: *"Kunskap och betyg"*, Studentlitteratur, Lund, 2003

Mejding, Jan (red): *"PISA 2003 – Danske unge i en international sammenligning"*, DPUs forlag, København, 2004

Miller, Tanja: *"Karaktergivning i praksis – 13-skalaen og gymnasiet"*, Dansk Institut for Gymnasiepædagogik, Syddansk Universitet, Ph.d. afhandling, Juni 2004

Reusch, Simon: *"2004-karakterer ved folkeskolens afgangsprøve i 9. klasse"*, http://www.uvm.dk/statistik/documents/karakterer2004fsa9-klv2_000.pdf

Skolverket: *"PISA 2000- Svenska femtonåringars läsförmåga och kunnande i matematik och naturvetenskap i ett internationellt perspektiv"*, Skolverket, Stockholm, 2001.

PISA Copenhagen 2004

– The Competence of 9th Form Students in Copenhagen

Niels Egelund and Beatrice Schindler Rangvid

Abstract

The PISA Copenhagen 2004 study is the first time PISA has been used in a municipal school development project. Thus PISA provides the baseline data for the project and in 2007 it will again provide the follow-up data for an evaluation of the 'effect' of the project. The 2004 study shows that it is possible to gain insight into important aspects of school functioning in a local area. Use of the test material from a previous PISA 2000 cycle makes valuable comparisons relatively easy. The main finding that school performances are heavily influenced by the social background and immigrant status of families living in the school district has important implications for school policy. Another point for consideration at a political level is the segregational effect of free schools. Danish free schools have attracted students with relatively good social backgrounds and academic records, students who would enhance the academic and social level in public schools via their peer effect.

Nordic abstract

PISA-København gennemført I 2004 repræsenterer første gang, hvor PISA er blevet brugt I et kommunalt skoleudviklingsprojekt. På den måde giver PISA data om det niveau, kommunens elever ligger på ved starten af projektet, og PISA vil igen i 2007 give oplysninger om niveauet, når projektet sluttes. PISA-København data fra 2004 viser, at det er muligt at få indsigt i mange vigtige aspekter af skolernes funktion i et lokalområde, og brugen af PISA-testen fra 2000 giver værdifulde sammenlignings-muligheder. Det væsentligste fund er, at elevresultater er stærkt sammenhængende sociale baggrund og forekomst af tosprogethed i et skoledistrikt, hvad der har vigtige skolepolitiske implikationer. Et andet væsentligt punkt til overvejelse er den segregerende effekt, de frie skoler har. Danske frie skoler har tiltrukket elever med relativ god social baggrund og faglige resultater, elever som gennem kammerateffekten kunne have bidraget positivt til det faglige og sociale niveau i de offentlige skoler.

In 2003 the local government of Copenhagen decided to carry out a PISA study in Copenhagen schools. The test, conducted in spring 2004, was a replication of the PISA 2000 study (Andersen et al. 2001) with a few extra items included in the student questionnaire. In the study all 59 municipal primary and lower-secondary schools with 9th form classes participated, as did 24 so-called 'free schools' (private independent schools). Seventeen schools free schools did not wish to participate. The total number of schools participating in the study is 83, comprising 2352 students out of a possible total of 2740. For various reasons, 14% did not participate – with a highly significant difference in dropout rate at the various schools (0-52%). In contrast to the international PISA study, all the students who participated in PISA Copenhagen were enrolled in the 9th form at the time the study was conducted. The data collection in the spring 2004 was the first phase in a school development programme,; further data from reading tests, interviews and observations will be used to change practice, followed by a second phase of data collection (again a replication of the PISA 2000 study) in 2007.

Seen from a Nordic perspective, the PISA Copenhagen study has interesting aspects: 1) It shows how PISA instruments and procedures can be used in a municipal school development programme. 2) It is the first time in the Nordic countries where the number of bilingual students in the study makes a detailed analysis of this group possible.

The results

The results from PISA Copenhagen 2004 are reported in Egelund & Rangvid (2005). The main results from the tests in reading, mathematics and science are summarised in table 1. Here it can be seen, that in PISA Copenhagen, the results are approximately 20 scale points below the average for Denmark as a whole. This placement can be linked to the fact that the number of bilingual students in Copenhagen is greater than the national average. When the bilingual students' results are omitted, Copenhagen is above the national average, except in mathematics, when Copenhagen is slightly below the average. For reading and science the differences are barely significant. The bilingual students in Copenhagen are approximately 100 scale points below the Danish students on the PISA scale, which means that only 15% of the Danish students are at a level that is below the average for bilingual students. The free schools are approximately 50 scale points above the municipal schools, but the free schools represent a very diverse group. The level of performance at Danish free schools is very high, while at ethnic free schools it is very low, with a spread of between 100 and 119 scale points.

   The share of students without functional reading skills in the PISA 2000 study was 18%. In Copenhagen, it is 24%, with 14% of Danish students and 51% of

Table 1 *Distribution of average scores on the PISA scale*. Top row shows PISA 2000 results from Denmark as a whole. Following rows are from PISA Copenhagen. First section for all students, second for Danish students, third for bilingual students*

	Reading Total	Reading Information	Reading Interpretation	Reading Reflection	Mathematics	Science
PISA2000	497	498	496	502	514	481
Cph total	478	471	480	481	478	461
Girls	488	477	489	498	470	453
Boys	469	468	473	466	489	471
Municip. sch.	465	458	467	466	468	451
Free sch.	516	508	518	523	506	492
Danish free	528	521	529	534	518	505
Ethnic free	428	414	433	443	411	386
Danish	507	501	509	509	508	487
Bi-lingual	414	405	417	419	413	396
Immigrants	404	394	408	406	404	387
Decendants	421	414	424	427	420	402
Public sch.	398	390	402	402	403	390
Free sch.	466	458	470	477	448	419

*Scores are derived from PISA scale Plausible Values

bilingual students showing poor reading skills. For first generation bilingual students, the share is 55%, while for second generation bilingual students it is 47%. The difference is statistically insignificant.

The differences among schools are very significant. Four schools in Copenhagen have an average for reading skills that is above Finland's national average in the PISA 2000 study, while 15 schools are below Brazil's national average (Brazil being the country with the worst performance in the PISA 2000 study). The greatest spread among schools in Copenhagen is in the science domain.

The same trend shown in Copenhagen towards a wide spread of achievement level among all students can be seen throughout Denmark as a whole – i.e. there are relatively many low-performing students, and this is the case for all three domains. The ethnic free schools in particular have a high number of weak students, although this should be seen in the light of the students' socio-economic backgrounds.

School demographic factors

There is no correlation between academic test results and classical school demographic factors such as school size or class size. The study also considers whether there is a correlation between school size and class size and some of the central social and welfare-related factors. This does not appear to be the case either.

Another important point of interest is the ratio of Danish versus bilingual students in the classes and schools in Copenhagen. The study shows that there is a relation between the ratio and performance; bilingual students perform poorly in schools with a high concentration of bilingual students. Moreover there is a significant general peer effect, whereby proficient students are able to raise the performance of weaker students in a class.

Benchmarking of schools

To achieve a fairer basis for comparison than is possible with the 'raw' data from the academic tests, which to a high degree reflect the socio-economic backgrounds of the students, a correction for social background was made using statistical methodology (Raudenbusch & Willms, 1995). While this kind of correction is controversial, it actually provides a superior point of reference for comparison – although there will still be a large number of unknown factors that are not accounted for. In this connection, one factor that is problematic is that some schools do not have many 9th form students and another is that the participation rate differs from school to school. An attempt has been made to account for the first factor, while the impact of the latter is not known. The findings must, therefore, only be viewed as a forerunner to further research.

The results suggest that student performance at some schools was more than 60 scale points higher than expected, while student performance at other schools was about 50 scale points lower than predicted by the statistical model.

We also made an attempt to identify schools which are particularly successful at boosting the achievement levels of students with weak socio-economic or immigrant backgrounds. However, the dataset is too small for rigorous analysis along these lines.

Student well-being and social relations

With regard to student well-being and social relations, the findings from PISA Copenhagen do not differ significantly from the findings for Denmark as a whole in the PISA 2000 study. Only 6% of students report *feeling like outsiders*. *Boredom* is not an uncommon phenomenon, but only about 15% indicate that they are often very bored. More than 10% of students indicate that there is a high degree of

disturbing noise and disruption, while just about 20% feel that there is only a moderate degree of disruption.

Although these findings resemble those for Denmark as a whole, it is particularly noticeable that students born outside Denmark are most likely to feel like outsiders – and that municipal school students are more likely to feel like outsiders than private school students. Also, schools with relatively poor academic performance have the highest number of students who feel like outsiders. The problem of boredom is greater among Danish students than bilingual students, and is least significant at ethnic free schools. Another rather curious finding is that there is a higher level of boredom at schools with good student performance. Noise and disruptions are a greater problem at municipal schools than at free schools, and reading performance is weakest at the schools with the highest number of students complaining about disruptions.

Another characteristic is that there is an extraordinary difference in Copenhagen from school to school in the findings for all three well-being areas, which again suggests that there is considerable potential for improvement.

Relationships with teachers

The students in Copenhagen reported on their relationship with their teachers, and in this area the distribution of the answers is the same as in PISA 2000. Approximately 30% do not feel that they have a good relationship with their teachers, and approximately 20% indicate that their teachers are not interested in them. There is an overrepresentation of bilingual students in the groups which indicate a poor student-teacher relationship. The problem is most significant at municipal schools and least significant at the Danish free schools, and there is a correlation between a good relationship to teachers and good reading performance.

Out of school activities

In the area of out of school activities, the study examines three key activities where the students in Copenhagen differ from other Danish students.

The amount of homework in the subject of Danish usually ranges between 1 and 3 hours a week, and girls spend noticeably more time on homework. Bilingual students spend more time on homework than Danish students, and the highest amount of homework is given at the ethnic free schools. The weakest students spend, by far, the least amount of time on homework, while the strongest students spend an average amount of time on homework.

Approximately a quarter of students do not voluntarily read in their spare time, and there are twice as many boys as girls in this group. Bilingual students read more than Danish-speaking students, and students from ethnic free schools read

the most. There is no correlation with reading competency in PISA beyond the fact that students who never read perform significantly less well than other students.

Around half of the students do not have any desire to visit a bookstore or library. In this area there is also generally a significant difference between boys and girls, with girls more likely to indicate an interest. Likewise, bilingual students and students at ethnic free schools also have a more positive attitude. There is a strong statistical correlation between a positive attitude towards bookstores and libraries and reading performance.

Personal and social competencies

There are no differences in the average results for items testing personal and social competencies, as investigated in the PISA study, between students in Copenhagen and students in Denmark as a whole.

With regard to learning strategies, it appears that students in the ethnic free schools in particular show a relatively high use of memorisation, the ability to relate tasks to known and relevant knowledge and the ability to control the learning process. The ability to relate to known and relevant knowledge is also, to some extent, a characteristic of the Danish free schools. All three learning strategies have a positive correlation with good reading performance. Two of these learning strategies, use of memorisation and control over the learning process, would traditionally be considered 'old-fashioned'.

At the ethnic free schools, the influence of career-related motivational factors is stronger than at the other types of school, just as students at these schools show greater perseverance. Students at the ethnic free schools also show the highest level of interest in competition, students at Danish free schools show the lowest level of motivation for collaborative work. There are no clear correlations between the motivational factors studied and reading performance.

Students at the free schools show most self-control, while there are no differences in the students' self-confidence and self-perception among the three school types. There is a very clear correlation between students with a positive sense of self-control and high self-confidence and good reading performance. In the case of self-perception, however, there is a negative correlation, i.e. that poor readers consider themselves to be better students than they are and vice-versa, which is difficult to understand.

ICT

Approximately 80% of students in Copenhagen have a computer they can use at home, and this is especially the case for students, the majority of whom are from the Danish free schools. As regards access to computers, students at the ethnic free

schools have less access than students at the municipal and Danish free schools. Around 70% of students feel that they are comfortable using a computer to write school assignments, but here again students at the ethnic free schools are at a disadvantage. Access to computers at home and feeling comfortable with using them for school assignments correlates with good reading performance, while access to a computer at school is less significant.

Between 60 and 70% of students use computers at home almost every day, with a slightly higher percentage being at free schools than at municipal schools. On the other hand, the use of computers at school is highest at municipal schools and lowest at ethnic free schools. There are relatively few students who use computers at the library. However, the student groups who are least likely to have access to computers at home are most likely to use library computers.

Between 60 and 70% of students also use the Internet and e-mail/Internet chat rooms every day, and the frequency is highest for ethnic free school students and lowest for municipal school students. There is a positive correlation between good reading performance and use of computers at home, and a negative correlation with use of computers at school and, especially, at the library. This is clearly related to the fact that students from the strongest social backgrounds are more likely to have access to computers at home, while students from the weakest social backgrounds are more likely to use the library.

Approximately 25% of students play computer games almost every day. Municipal school students spend most time on this activity, while students from the ethnic free schools spend least time. Word processing is used more often than computer games, and most often for students at the Danish free schools. Spreadsheets, drawing and graphics programs are used relatively seldom and are least likely to be used by municipal school students. Students at ethnic free schools are most likely to use educational software.

Between 60 and 80% of students find it important and interesting to work on a computer. Municipal school students have the most positive attitude, while students from the ethnic free schools are least positive. Interestingly, it appears that the students with the weakest reading competencies are most likely to have a positive attitude towards ICT. It also appears that there is only a weak correlation between attitude towards and use of ICT and parents' educational background, while there is a very significant gender difference in attitudes, with boys more heavily represented among students with positive attitudes towards ICT.

Information from school principals

School principals have provided a good deal of information on the conditions at their schools regarding administrative and demographic factors, as well as attitudes to students and teachers.

The likelihood that students are moved to other schools applies to both municipal and free schools, although the reasons often differ. The likelihood that students are moved from free schools because of poor academic performance or behavioural problems is greater than for municipal schools, especially at high-performance schools. And these moves are not necessarily initiated by the parents. However, the need for special educational assistance does not appear to be the main reason for moving students.

Sub-standard school building conditions and inadequate space are considered major problems by school principals, especially at municipal schools. Similarly, inadequate provision of teaching materials and computers as well as sub-standard heating, ventilation and lighting systems are also problems at municipal schools. The free schools have the poorest selection of teaching material in their libraries. However, there is no clear correlation with poor reading performance at the schools.

Teacher-related conditions have, for the most part, been studied with regard to their influence on reading performance. From this perspective, it appears that factors such as low expectations of students, poor relationships between students and teachers, inadequate appreciation of academic skills and a high degree of teacher turnover correlate with poor reading performance. Low expectations are most likely at municipal schools and ethnic free schools. Appreciation of academic skills is lowest at ethnic free schools and highest at Danish free schools. Teacher turnover is the greatest problem at free schools.

With regard to the use of evaluations, it is not possible to conclude that the use of standardised tests correlates with reading performance, while there is a positive correlation with tests prepared by teachers and general use of teacher evaluations.

Student-related conditions have a significant correlation with reading performance. Statistically significant factors include socially weak backgrounds, a tendency to disrupt classes, lack of respect for teachers, absence from school and lack of help with homework from parents. Not enough class time and abuse of alcohol and/or drugs by students are not significant factors.

A profile of bilingual students

Of the bilingual students who participated in the study, seven out of eight speak a non-Western language, the most common language being Arabic. A higher proportion of bilingual students than Danish students live with both their parents, and the number of siblings is greater. Eight per cent of bilingual students come from families where neither parent has an education, and almost 40% come from families where neither parent has a full-time job. Even when the socio-economic status of the parents is taken into account, there are fewer educational resources available in the homes of bilingual students.

Reading habits and leisure time activities vary somewhat for the two groups as bilingual girls read more than both Danish students and bilingual boys. Participation in (Western) cultural activities is somewhat greater among Danish students than among bilingual young people. With regard to ICT use and skills, the differences between bilingual and Danish young people are less than for most of the other factors considered in the study. And in some areas, such as the use of educational software and the Internet, bilingual girls, in particular, are more active than Danish girls.

A profile of students with weak socio-economic backgrounds

A comparison between students with the weakest and the strongest socio-economic backgrounds (the weakest and strongest 15%) provides many illustrative differences. First and foremost, there is a strong overrepresentation of bilingual students among those with weak socio-economic background and a certain prevalence of students from single parent families. Labour market attachment is also extremely different, as are cultural interests at home, which in turn affect the help available for homework. Children from weaker socio-economic backgrounds are forced to seek help elsewhere, e.g. from siblings, friends and tutors at school.

Reading habits and ICT use vary relatively little between the two groups, although girls from socially disadvantaged families spend a significantly greater amount of time reading for their own pleasure than both boys and girls from socially strong families. With regard to well-being at school, the differences are minimal, with one exception: the most socially disadvantaged students experience boredom less often than the more well-off.

The greatest differences in leisure time activities relate to participation in classical cultural activities like opera, ballet and theatre, as well as all types of concerts.

Conclusion

The PISA Copenhagen study shows that it is possible to gain insight into important aspects of school functioning in a local area. Use of the test material from a previous PISA cycle makes valuable comparisons relatively easy.

The main finding that school performances are heavily influenced by the social background and immigrant status of families living in the school district has important implications for school policy. Another point for consideration at a political level is the segregational effect of free schools. Danish free schools have attracted students with relatively good social backgrounds and academic records, students who would enhance the academic and social level in public schools via their peer effect.

The results from PISA Copenhagen 2004 are being used in a school development project: researchers are looking at the daily functioning of the schools and giving advice to school management and teaching personnel. In 2007 the test will be repeated to see if the project has been fruitful.

References

Andersen, A.M., N. Egelund, T.P. Jensen, M. Krone, L. Lindenskov & J. Mejding (2001): *Forventninger og færdigheder – danske unge i en international sammenligning.* København. Socialforskningsinstituttet. (In Danish)

Egelund, N. & Rangvid, B.S. (2005): *PISA-København.* København. Amternes og kommunernes forskningsinstituts forlag. (In Danish)

Raudenbusch, S.W. & J.D. Willms (1995): The estimation of school effects. *Journal of Educational and Behavioral Statistics.* 20, 307-336.

Chapter 17

Participants in PISA 2000 – Four Years Later

Torben Pilegaard Jensen and Dines Andersen

Abstract

This article presents the preliminary results of the newly established PISA Longitudinal database which offers insights into aspects such as how social backgrounds, attitudes and life values affect young people's choice or rejection of education and their subsequent status in the labour market. The long-term research perspective of the project is to illuminate the path from childhood through adolescence and education to adulthood, rooted as it is in family and working life.

Almost all young people – including those with poor reading skills – begin a post-compulsory course of education sometime after completing basic schooling. However, compared to young people with better reading skills those with poor reading skills show a significantly greater risk of not commencing or completing such a course of education.

Four years after completing basic compulsory education, the educational status of the young people is primarily determined by the reading skills and academic self-image they possessed in the 9th form. Social background has an indirect impact on reading skills, but it also has a direct impact on the young people's status in the educational system. Furthermore, gender and orientation towards career or practical work are factors which influence independently what kind of post-compulsory education the young people choose.

Nordic abstract

I denne artikel fremlægges de første resultater fra den nyetablerede database PISA Longitudinal, der bl.a. kan belyse, hvilken rolle de unges sociale baggrund, deres faglige færdigheder i grundskolen og deres holdninger og livsværdier spiller for deres valg og fravalg af uddannelse og senere placering på arbejdsmarkedet. Vejen fra barndom over ungdom og uddannelse til voksen med forankring i familie og arbejdsliv er forsknings-perspektivet på længere sigt.

Næsten alle unge – også dem med ringe læsefærdigheder – starter i forlængelse af grundskolen eller lidt senere på en ungdomsuddannelse. Men unge med ringe læse-færdigheder har dog en væsentlig større risiko for ikke at starte eller ikke at gennemføre uddannelsen sammenlignet med unge med bedre læsefærdigheder.

4 år efter grundskolen er de unges uddannelsesmæssige position primært bestemt af deres læsefærdigheder og faglige selvopfattelse i 9. klasse. Den sociale baggrund spiller indirekte ind i forhold til læsefærdigheder, men har derudover en direkte effekt på de unges placering i uddannelsessystemet. Endvidere spiller kønnet og de unges orientering mod karriere eller praktisk arbejde en selvstændig rolle for, hvilken type ungdomsuddannelse de unge vælger.

Introduction

The PISA studies have provided the participating countries with a tool for evaluating, in an international context, national efforts in the area of basic compulsory education. The stated objective of PISA is to measure the competencies of the pupils in relation to the demands they will face 'in the real world' after basic compulsory education, where the ability to acquire new knowledge, among other things, is the principal focus. These demands, which are set at the international level in the OECD, are not intended to be identical with the national curricula. In general terms, PISA is, thus, primarily an evaluation of the basic compulsory education system's ability to give the pupils competencies that will benefit them after basic compulsory education.

Questions have been raised from several quarters as to whether the PISA measurements in broad terms provide relevant information on the qualifications of young people. Some say that by focusing on skills in reading, mathematics and science, many other skills are disregarded, skills which are important in the globalised world the young people are supposed to be prepared for. In this regard, it should be noted that PISA contains many other more socially oriented indicators of the skills and competencies of the participants. The studies[1] upon which this article draws are based on interviews with young people 4 years after leaving school and, thus, present a longitudinal perspective. This means that reading skills constitute just one of several competencies that help to explain whether the young people complete an education after they finish basic compulsory education. One of the requirements of future society is the ability to change. Therefore, modern people must be prepared to change, be ready to acquire new knowledge.

---

1. See Andersen, 2005.

Lifelong learning is the catchphrase of the day. One of the basic prerequisites for being able to acquire new knowledge is to have adequate reading skills. This proposition appears to state the obvious, but it should nevertheless be tested. Is it correct that people with poor reading skills have only limited opportunities to advance in the educational system? Not all knowledge in life comes from the world of books. Is not the only sure test of the argument to follow a group of young people after they leave compulsory education and observe how things work out for them later on? Who obtains further education and who does not?

PISA Longitudinal

The alliance[2] behind the Danish component of PISA 2000 was quick to see the potential of a possible follow-up study, based partly on data from coupling the database to registers at Statistics Denmark and partly on follow-up interviews. The complete database, which will be regularly expanded with additional data from Statistics Denmark's administrative registers and from new rounds of interviews, is called *PISA Longitudinal*. In 2005 it contained data from the Danish PISA 2000 assessments (about 4,000 young people born in 1984) and data from follow-up interviews in 2004 of about 3,100 of these young people.

By following these young people for a number of years PISA Longitudinal gives us the opportunity to see what is actually happening with every single individual. Which conditions are of crucial importance for young people to get an education and a permanent position in the labour market? Is it true that the measure of reading skills obtained in PISA is a good indicator for how well young people will do later on? By expanding the Danish PISA study into a longitudinal study the measurement of reading skills gets validated compared to reality.

Denmark is one of the few countries participating in PISA 2000 that has been able to expand and upgrade the original data set. Canada and Switzerland are in the process of establishing similar data sets. This means that PISA Longitudinal, which will be accessible to all countries[3], will play a central role in an international research context.

---

2. PISA is carried out in Denmark as a collaboration between akf, institute of local government studies – Denmark, the Danish National Institute of Social Research and the Danish University of Education.
3. An English-language version of the database will become available on a website currently under construction.

Post-compulsory education in Denmark

The majority of 15-year-olds who participated in PISA 2000 were in the 9th form, which is the highest basic compulsory form in Denmark. Before the young people move on to a post-compulsory course of education, they may choose to take an elective 10th school year. A relatively large number choose this option (approx. 60%).

Post-compulsory education has two main paths: 1) *Upper secondary education* gives young people academic competencies (for courses in higher education) and takes three years. This path is chosen by roughly half of all students; 2) *Vocational education* gives young people vocational competencies and have varying durations, depending on their focus, but approximately three years is most common. Most vocational courses of education are begun after completing the elective 10th school year.

A discussion is currently taking place in Denmark about whether the 10th school year has become an unnecessary detour in the long course of education which only postpones the time when young people finish their education. Ongoing deliberations about the future status of the 10th school year should be viewed in the light of the dual purpose it serves: 1) to help those young people who are *not sufficiently ready* (mature) to choose a post-compulsory course of education that is right for them, considering their interests and skills, and 2) to give them the opportunity *to improve their academic skills* before beginning a post-compulsory course of education that might otherwise be more than a somewhat weak pupil can handle.

Because of this duality, it is difficult for outsiders to determine whether the 10th form serves its intended purpose. It is an option not only for those with weak academic skills, but also for young people who do not know what they want to do with their lives. However, from an overall perspective, this should not prevent us from asking whether basic compulsory education in general lives up to the expectation of educating young people so that they are able, after the 9th form, to move on to a post-compulsory course of education, since more than half a generation of young people feel they need the extra 10th school year.

Our analysis of which young people chose a 10th school year before moving on to a post-compulsory course of education shows that there is a clear correlation with the individual pupil's reading skills.[4] Whereas two out of three pupils with reading

---

4. The study utilises the aggregate measurement for reading skills, the WLEREAD variable.

skills at the lowest levels (level 1 or lower) preferred to attend the 10th form, the respective figure among pupils with the best reading skills (level 5) was only 40%. One likely reason why the correlation is not stronger relates to the second purpose of the 10th form – to give more time for young people to work out what they want to do.

Taking stock – four years later

The follow-up interviews took place four years later in the spring of 2004. For many of the participants, this was immediately prior to completing their post-compulsory course of education. In other words, these young people were *still studying*. About one in three had already *completed* their post-compulsory course of education and had either *begun a new course of (higher) education* (approx. 12%) or had (temporarily) *dropped out of the educational sector* (approx. 20%), obtaining instead vocational work or taking time off. Initially, one might think that those who had left the educational sector were graduates with a diploma from a vocational programme, but this was not the case. By far the majority of this group had an upper-secondary school education, and so they most likely considered their situation to be temporary (a sabbatical year to experience the world and find out which course of higher education was right for them). Finally, one in ten had neither completed a course of education beyond basic compulsory school nor was in the process of doing so. This is called the *residual group* without a qualifying education. The majority of the residual group have at one time or another attempted to obtain a post-compulsory education, but they have not been able to follow through and have dropped out again.

How do the four groups of young people differ? Why have some gone further in their studies than others, and why have some made no progress? Answers to such questions can be found through statistical analysis of the educational status of the young people at the age of 19. A wide variety of conditions that describe the young people's family backgrounds, cultural transfer within the family (through interviews with their parents about culture and politics as well as their own high-culture activities), their interests, relationships with teachers and classmates in the 9th form and their academic level have been taken into consideration in the study, which shows that a number of these conditions are significant – even after correction for social background. The strongest of these conditions are two indicators relating to the academic level of the young people: their *test results for reading proficiency* and their *self-assessment of academic level* in the 9th form (that they perceive themselves as good at most subjects). The better the reading skills, the greater the likelihood that the young person will have completed a course of education; whereas the lower the score on the reading test, the greater the likelihood that the young person will

become part of the residual group. The variation with regard to self-assessment of academic level basically follows the same pattern. This is because there is a strong correlation between test results and self-assessment of academic level. Young people who, in objective terms, do well in a particular subject will also see themselves subjectively as high achievers. But this self-knowledge also has an independent value as an explanatory factor.

Other aspects also play a significant role, for instance *gender* (currently girls are more likely than boys to choose an upper secondary education and, therefore, are more likely to have completed a post-compulsory course of education). Furthermore, *family background* plays a crucial role, something our study appears to underestimate. This is because family background has both a direct and an indirect (i.e. hidden) effect. The latter consists of the fact that family background greatly influences the development of the young person's reading skills and self-image. Since these aspects are independent factors in the study, it is their effect that is measured. However, in addition to the hidden effects of the young people's family backgrounds, the study also sheds light on the direct effects[5].

Difficult, but not impossible

There is nothing new in the argument that reading skills at a certain level are considered a key prerequisite for successfully obtaining an education. Based on the test results in PISA 2000, it was determined that 'at least 18% of pupils in Denmark are expected to have difficulties utilising reading to acquire new knowledge' (Andersen et al. 2001). This is the group of young people sometimes referred to in Denmark as 'functional illiterates' who can be expected to be at particularly great risk of ending up in the residual group without an education. Our studies of how things actually turned out for the young people during the first four years after completing their basic compulsory education have produced several results. First, the expectations appear to have been fulfilled. Whereas just 4% of the very strong readers were part of the residual group four years later, the proportion of young people in that group with very poor reading skills was 17%. As mentioned above, the residual group comprises young people who have not begun a course of education as well as young people who have begun one and dropped out again. We may not have seen the last of the dropouts. In which case, who is at greatest risk of dropping out? A good guess would be young people with poor reading skills. At least we can see that the reading skills among those young people who have attempted to obtain a post-compulsory education but dropped out during the process are significantly weaker than among those young people who are still working on their education.

The second significant result is that the majority of the very weak readers were still in the process of obtaining an education and that one in five young people with level 1 reading skills and one in eight with readings skills lower than level 1 had, in fact, already completed a course of education. This clearly indicates that the correlation between reading skills and advancement in the educational system should be understood as the better a person's reading skills, the greater the *likelihood* that that person will complete a post-compulsory formal education. However, it is *by no means a law of nature* that people with poor reading skills are unable to complete a course of education, nor is it a fact that people with excellent reading skills will also obtain a formal qualifying education. It is all a question of greater or lesser probabilities (chances).

As a natural extension of this, one might ask what types of education these young people with poor reading skills have completed. While this still needs to be determined, it is presumed that they are likely to be educational programmes with a high degree of practice-oriented content and, perhaps, of relatively short duration.

The study also suggests that young people with good reading skills come through the educational system at a faster rate than weak readers, and that more good readers proceed to a higher education. In other words: poor reading skills appear to extend the duration of a course of education up to a given point. The majority of weak readers had either not begun a course of education or were still working on their education. Among the strong readers, half had already completed their post-compulsory course of education and had entered the labour market or had moved on to a new (higher) education.

This suggests that in the long term, there would be several advantages to be gained from improving reading skills (especially for the weakest readers). Partly, this would give more young people the opportunity to complete a post-compulsory education and, partly, it would enable the education sector to increase the rate of completion, for which there seems to be a great need in Denmark, where young people generally spend more time in the educational system than they do in most of the other countries with which Denmark is usually compared.

Profiles of young people following different courses of education

Through interviews at the age of 19, the young people indicated how much emphasis they placed on various reasons for their choice of activity after basic compulsory education. Some stemmed from a desire to do *something else* rather

than go to school. As expected, young people from the residual group, in particular, placed great emphasis on the fact that they were tired of going to school; they wanted to earn money or gain vocational experience. For all other young people, these arguments played a very small role.

Those young people who embarked on a course of education expressed a variety of motives for choosing a particular education. The strongest reasons involved long-term planning, for instance consideration of *future employment options*, the young people's *academic interests* and their expectations that their choice will lead to a *good working environment*. Some emphasis, though not as much, was also placed on the prospect of a job with *high status and a high income*. This average portrait of the entire group is also typical of the majority of young people in general upper secondary school in Denmark. In general, the young people who chose a *commercial upper secondary school* prioritised in the same way as those in general upper secondary school. However, they appear to have placed more emphasis on getting out of the general school environment, and on the prospect of status and income. The young people who chose a *technical upper secondary school* differed greatly from the other upper secondary school paths in that they placed a good deal of emphasis on their academic interests.

The other main post-compulsory education tracks are the *vocationally oriented programmes* at business, technical, social and health service colleges. Young people who chose these programmes prioritised their reasons slightly differently. Overall, they placed more emphasis on being tired of school and on positive experiences in vocational traineeships, while the prospect of high status and income clearly played a less significant role.

Of course the young people's motives and interests do not come out of nowhere. They are a result of influences from home, friends and school. The academic ballast and self-knowledge that young people acquire are especially important for their personal development. The young people who emphasised their *future employment prospects*, getting a *job with high status* and their *academic interests* also scored much higher on most indicators than young people who did not consider these aspects important. Academically capable young people with a strong self-image, positive experiences from school, strong family backgrounds and a high degree of cultural transfer from their parents placed the greatest importance on these aspects – these are young people who are expected to do well in life. The opposite is the case for young people who emphasised being *tired of school* and a desire to *earn money*.

The young people also indicated the importance of working within a specific field in their future working life. These wishes combined with the reasons for choice of

course of education were analysed with a view to revealing areas where there is so much correlation in the responses that it is most probably attributable to one underlying aspect, which has carried through to several of the questions. One such factor analysis led to three overriding motives: an inclination towards *practical work*, *career-orientation* and a dependence on *support or other encouragement from their surroundings*.

After pinpointing the three overriding motives for choice of education, they were included in a complete analysis of the backgrounds and motives for the young people's choice of three very different educational tracks: technical college and commercial and general upper secondary school. The technical colleges offer a wide variety of programmes within traditional trades (mechanics, iron and steel, building, foods) as well as a variety of service-oriented disciplines. It would undoubtedly make sense to differentiate between the individual areas, which are distinguished by having very different gender recruitment bases. But here they are treated as a single group of educational programmes. When the significance of social background is compensated for, the analysis shows that choices among the three main tracks are based on many conditions.

Overall, young people at technical college primarily differ from the comparison group (young people at general upper secondary school) in that they are significantly more oriented towards practical work and have attained a lower academic level. Career orientation and dependence on the acceptance/support of their surroundings are also less significant. The same applies to the educational level of their fathers (more so in the residual group). Other things being equal, cultural transfer within the family and experiences at school do not play a significant role. Because there is presumably a correlation between an inclination towards practical work and low academic qualifications (which, to some extent, discourage the young people from choosing an academic upper secondary education), the conclusion must be that what distinguishes young people taking a technical-vocational education from those taking a general upper secondary education are the academic qualifications needed for completing such an education.

The comparison between young people who chose a commercial upper secondary education and those taking a general upper secondary education looked a little different. Initially, they did not differ very much with regard to inclination towards practical work. Their academic level was lower, and they had more negative experiences with school than those who chose a general upper secondary education. On the other hand, they had a stronger career orientation. Their parents' vocational and educational levels were the same, but they experienced less cultural transfer.

Thus, the study suggests that general upper secondary education is considered, to some extent, to be a continuation of basic compulsory education and, consequently, appeals to those young people who felt comfortable in that environment. General upper secondary education also upholds the classical ideal of education, which is why it appeals to young people from homes with strong cultural transfer. On the other hand, commercial upper secondary education seems to appeal to those young people who have the skills for an upper secondary education, but who have had negative experiences with schoolmates and teachers in basic compulsory school – that is, young people who need a change of pace. They are more focused on having a career and have not been 'overloaded' with high-cultural transfer within the family. With courses of education at the same academic level, it is to be expected that the other more cultural factors, rather than the academic factors, play a decisive role in choice of education.

A Nordic perspective

This article shows how the data from PISA can be used as part of a longitudinal analysis supplemented with data collected at a later date. Because access to the original data from PISA 2000 is free, it was in fact a relatively manageable task to create the PISA Longitudinal database. The database is Danish, but it is based on the international PISA study, which means that if a number of countries had decided to become involved, it could easily have been established as a cross-national study with coordinated data collection. PISA continues with new rounds of data collection and new opportunities to establish a cross-national study. In the meantime, the results of the Danish PISA Longitudinal data could be useful in other Nordic countries as they are quite similar in culture and social structure. Thus, when differences in educational possibilities are taken into account, the results indicating the competencies that play a central role in young people's educational choices after compulsory school in Denmark are likely to be realistically relevant to young people in the other Nordic countries.
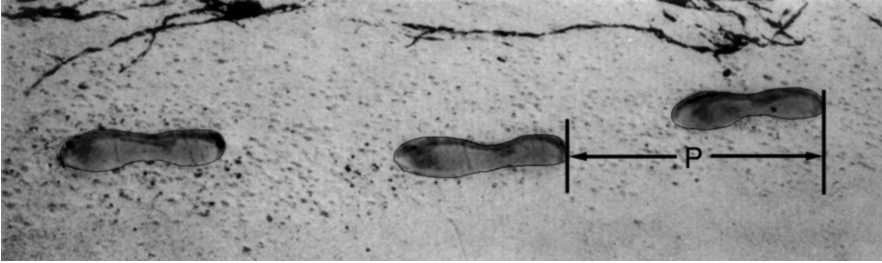
References

Andersen, Dines (2005): *4 år efter grundskolen*. [4 years after basic compulsory education]. akf forlaget, Copenhagen. See English summary on www.akf.dk

Jensen, Ulla Højmark and Torben Pilegaard Jensen (2005): Unge uden uddannelse [Young people without an education]. SFI, Copenhagen. See English summary on www.sfi.dk

Appendix

WALKING

The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

For men, the formula, $\dfrac{n}{P} = 140$, gives an approximate relationship between n and P where,

> n = number of steps per minute, and

> P = pacelength in metres.

WALKING

M124Q03- 00 11 21 22 23 24 31 99

Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking.

Calculate Bernard's walking speed in metres per minute and in kilometres per hour. Show your working out.

WALKING SCORING 3

Full Credit

Code 31: Correct answers (unit not required) for both metres/minute and km/hour:
   n = 140 x .80 = 112.
   Per minute he walks 112 x .80 metres = 89.6 metres.
   His speed is 89.6 metres per minute.
   So his speed is 5.38 or 5.4 km/hr.

   Code 31 as long as both correct answers are given (89.6 and 5.4), whether working out is shown or not. Note that errors due to rounding are acceptable. For example, 90 metres per minute and 5.3 km/hr (89 X 60) are acceptable.
   • 89.6, 5.4.
   • 90, 5.376 km/h.
   • 89.8, 5376 m/hour [note that if the second answer is given without units, it should be coded as 22].

Partial Credit (2-point)

Code 21: As for code 31 but fails to multiply by 0.80 to convert from steps per minute to metres per minute. For example, his speed is 112 metres per minute and 6.72 km/hr.
   • 112, 6.72 km/h.

Code 22: The speed in metres per minute correct (89.6 metres per minute) but conversion to kilometres per hour incorrect or missing.
- 89.6 metres/minute, 8960 km/hr.
- 89.6, 5376.
- 89.6, 53.76.
- 89.6, 0.087 km/h.
- 89.6, 1.49 km/h.

Code 23: Correct method (explicitly shown) with minor calculation error(s) not covered by Code 21 and Code 22. No answers correct.
- n=140 x .8 = 1120; 1120 x 0.8 = 896. He walks 896 m/min, 53.76km/h.
- n=140 x .8 = 116; 116 x 0.8 =92.8. 92.8 m/min -> 5.57km/h.

Code 24: Only 5.4 km/hr is given, but not 89.6 metres/minute (intermediate calculations not shown).
- 5.4.
- 5.376 km/h.
- 5376 m/h.

Partial Credit (1-point)

Code 11: n = 140 x .80 = 112. No further working out is shown or incorrect working out from this point.
- 112.
- n=112, 0.112 km/h.
- n=112, 1120 km/h.
- 112 m/min, 504 km/h.

No Credit

Code 00: Other responses.

Code 99: Missing.

GROWING UP

YOUTH GROWS TALLER

In 1998 the average height of both young males and young females in the Netherlands is represented in this graph.
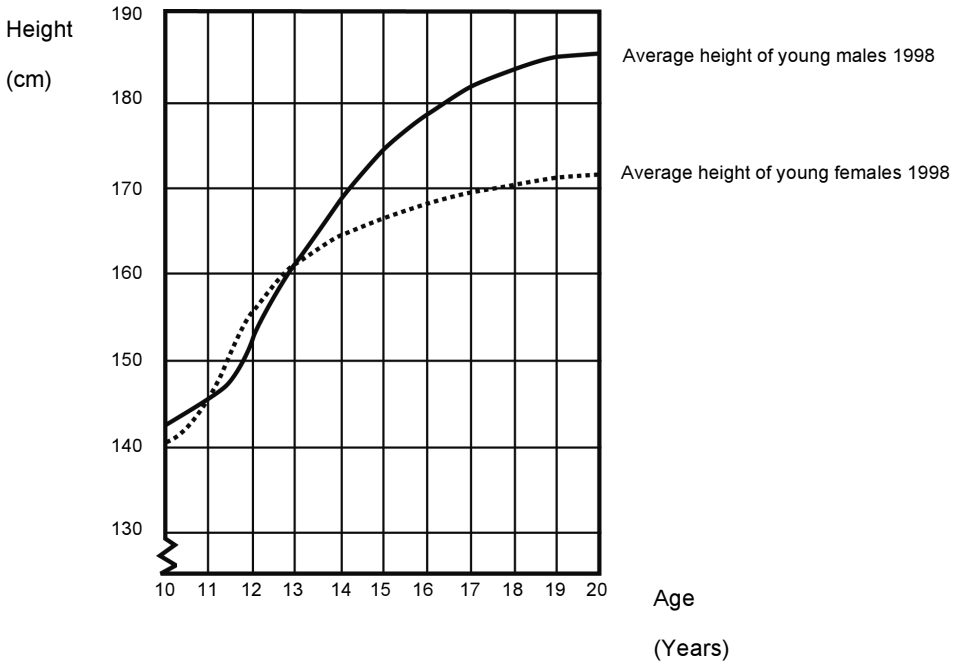
GROWING UP

Explain how the graph shows that on average the growth rate for girls slows down after 12 years of age. (Three lines for writing the answer)

GROWING UP SCORING 3

Full Credit
The key here is that the response should refer to the "change" of the gradient of the graph for female. This can be done explicitly or implicitly. Code 11 and code 12 are for explicitly mentioning about the steepness of the curve of the graph, while code 13 is for implicit comparison using the actual amount of growth before 12 years and after 12 years of age.

Code 11: Refers to the reduced steepness of the curve from 12 years onwards, using daily-life language, not mathematical language.
 • It does no longer go straight up, it straightens out.

- The curve levels off.
- It is more flat after 12.
- The line of the girls starts to even out and the boys line just gets bigger.
- It straightens out and the boys graph keeps rising.

Code 12: Refers to the reduced steepness of the curve from 12 years onwards, using mathematical language.
- You can see the gradient is less.
- The rate of change of the graph decreases from 12 years on.
- [*The student computed the angles of the curve with respect to the x-axis before and after 12 years*.]

In general, if words like "gradient", "slope", or "rate of change" are used, regard it as using mathematical language.

Code 13: Comparing actual growth (comparison can be implicit).
- From 10 to 12 the growth is about 15 cm, but from 12 to 20 the growth is only about 17 cm.
- The average growth rate from 10 to 12 is about 7.5 cm per year, but about 2 cm per year from 12 to 20 years.

No Credit

Code 01: Student indicates that female height drops below male height, but does NOT mention the steepness of the female graph or a comparison of the female growth rate before and after 12 years.
- The female line drops below the male line.

If the student mentions that the female graph becomes less steep, AS WELL AS the fact that the graph falls below the male graph, then full credit (Code 11, 12 or 13) should be given. We are not looking for a comparison between male and female graphs here, so ignore any reference on such a comparison, and make a judgement based on the rest of the response.

Code 02: Other incorrect responses. For example, the response does not refer to the characteristics of the graph, as the question clearly asks about how the GRAPH shows …
- Girls mature early.
- Because females go through puberty before males do and they get their growth spurt earlier.
- Girls don't grow much after 12. [*Gives a statement that girls' growth slows down after 12 years of age, and no reference to the graph is mentioned.*]

Code 99: Missing.

GROWING UP

M150Q02- 00 11 21 22 99

According to this graph, on average, during which period in their life are females taller than males of the same age? (Two lines for writing the answer.)

GROWING UP SCORING 2

Full Credit

Code 21: Gives the correct interval, from 11-13 years.
- Between age 11 and 13.
- From 11 years old to 13 years old, girls are taller than boys on average.
- 11-13.

Code 22: States that girls are taller than boys when they are 11 and 12 years old. (This answer is correct in daily-life language, because it means the interval from 11 to 13).
- Girls are taller than boys when they are 11 and 12 years old.
- 11 and 12 years old.

Partial Credit

Code 11: Other subsets of (11, 12, 13), not included in the full credit section.
- 12 to 13.
- 12.
- 13.
- 11.
- 11.2 to 12 .8.

No Credit

Code 00: Other responses.
- 1998.
- Girls are taller than boys when they're older than 13 years.
- Girls are taller than boys from 10 to 11.

Code 99: Missing.

ROBBERIES

ROBBERIES

A TV reporter showed this graph and said:

"The graph shows that there is a huge increase in the number of robberies from 1998 to 1999."
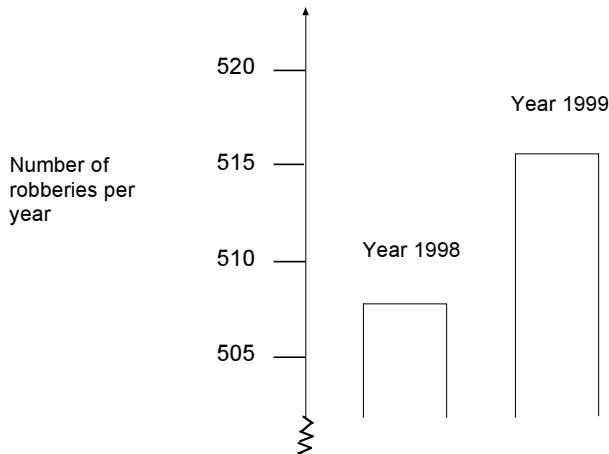
Do you consider the reporter's statement to be a reasonable interpretation of the graph? Give an explanation to support your answer.

ROBBERIES SCORING 1

[Note: The use of NO in these codes includes all statements indicating that the interpretation of the graph is NOT reasonable. YES includes all statements indicating that the interpretation is reasonable. Please assess whether the student's response indicates that the interpretation of the graph is reasonable or not reasonable, and do not simply take the words "YES" or "NO" as criteria for codes.]

Full Credit

Code 21: No, not reasonable. Focuses on the fact that only a **small part** of the graph is shown.
- Not reasonable. The entire graph should be displayed.
- I don't think it is a reasonable interpretation of the graph because if they were to show the whole graph you would see that there is only a slight increase in robberies.
- No, because he has used the top bit of the graph and if you looked at the whole graph from 0 – 520, it wouldn't have risen so much.
- No, because the graph makes it look like there's been a big increase but you look at the numbers and there's not much of an increase.

Code 22: No, not reasonable. Contains correct arguments in terms of ratio or percentage increase.
- No, not reasonable. 10 is not a huge increase compared to a total of 500.
- No, not reasonable. According to the percentage, the increase is only about 2%.

• No. 8 more robberies is 1.5% increase. Not much in my opinion!
• No, only 8 or 9 more for this year. Compared to 507, it is not a large number.

Code 23: Trend data is required before a judgement can be made.
    • We can't tell whether the increase is huge or not. If in 1997, the number of robberies is the same as in 1998, then we could say there is a huge increase in 1999.
    • There is no way of knowing what "huge" is because you need at least two changes to think one huge and one small.

Partial Credit

Code 11: No, not reasonable, but explanation lacks detail. Focuses ONLY on an increase given by the exact number of robberies, but does not compare with the total.
    • Not reasonable. It increased by about 10 robberies. The word "huge" does not explain the reality of the increased number of robberies. The increase was only about 10 and I wouldn't call that "huge".
    • From 508 to 515 is not a large increase.
    • No, because 8 or 9 is not a large amount.
    • Sort of. From 507 to 515 is an increase, but not huge.

[Note that as the scale on the graph is not that clear, accept between 5 and 15 for the increase of the exact number of robberies.]

Code 12: No, not reasonable, with correct method but with minor computational errors.
    • Correct method and conclusion but the percentage calculated is 0.03%.

No Credit

Code 01: No, with no, insufficient or incorrect explanation.
    • No, I don't agree.
    • The reporter should not have used the word "huge".
    • No, it's not reasonable. Reporters always like to exaggerate.

Code 02: Yes, focuses on the appearance of the graph and mentions that the number of robberies doubled.
    • Yes, the graph doubles its height.
    • Yes, the number of robberies has almost doubled.

Code 03: Yes, with no explanation, or explanations other than Code 02.

Code 04: Other responses.

Code 99: Missing.

EXCHANGE RATE

Mei-Ling from Singapore was preparing to go to South Africa for 3 months as an exchange student. She needed to change some Singapore dollars (SGD) into South African rand (ZAR).

EXCHANGE RATE

M413Q03 - 01 02 11 99

During these 3 months the exchange rate had changed from 4.2 to 4.0 ZAR per SGD.

Was it in Mei-Ling's favour that the exchange rate now was 4.0 ZAR instead of 4.2 ZAR, when she changed her South African rand back to Singapore dollars? Give an explanation to support your answer.

EXCHANGE RATE SCORING 3

Full Credit

Code 11: 'Yes', with adequate explanation.
- Yes, by the lower exchange rate (for 1 SGD) Mei-Ling will get more Singapore dollars for her South African rand.
- Yes, 4.2 ZAR for one dollar would have resulted in 929 ZAR. *[Note: student wrote ZAR instead of SGD, but clearly the correct calculation and comparison have been carried out and this error can be ignored]*
- Yes, because she received 4.2 ZAR for 1 SGD, and now she has to pay only 4.0 ZAR to get 1 SGD.
- Yes, because it is 0.2 ZAR cheaper for every SGD.
- Yes, because when you divide by 4.2 the outcome is smaller than when you divide by 4.
- Yes, it was in her favour because if it didn't go down she would have got about $50 less.

No Credit

Code 01: 'Yes', with no explanation or with inadequate explanation.
- Yes, a lower exchange rate is better.
- Yes it was in Mei-Ling's favour, because if the ZAR goes down, then she will have more money to exchange into SGD.
- Yes it was in Mei-Ling's favour.
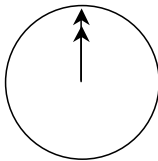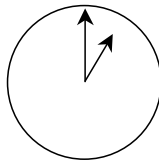
Code 02: Other responses.

Code 99: Missing.

INTERNET RELAY CHAT

Mark (from Sydney, Australia) and Hans (from Berlin, Germany) often communicate with each other using "chat" on the Internet. They have to log on to the Internet at the same time to be able to chat.
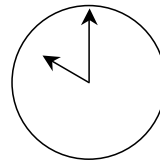
To find a suitable time to chat, Mark looked up a chart of world times and found the following:

| Greenwich 12 Midnight | Berlin 1:00 AM | Sydney 10:00 AM |

Question 1: INTERNET RELAY CHAT

M402Q01 - 0 1 9

At 7:00 PM in Sydney, what time is it in Berlin?

Answer: ....

INTERNET RELAY CHAT SCORING 1

Full Credit

Code 1:   10 AM or 10:00.

No Credit

Code 0:   Other responses.

Code 9:   Missing.

Question 2: INTERNET RELAY CHAT

Mark and Hans are not able to chat between 9:00 AM and 4:30 PM their local time, as they have to go to school. Also, from 11:00 PM till 7:00 AM their local time they won't be able to chat because they will be sleeping.

When would be a good time for Mark and Hans to chat? Write the local times in the table.

Place	Time
Sydney	
Berlin	

INTERNET RELAY CHAT SCORING 2

Full Credit

Code 1:   Any time or interval of time satisfying the 9 hours time difference and taken from one of these intervals:

Sydney: 4:30 PM – 6:00 PM; Berlin: 7:30 AM – 9:00 AM

OR

Sydney: 7:00 AM – 8:00 AM; Berlin: 10:00 PM – 11:00 PM
- Sydney 17:00, Berlin 8:00.

NOTE: If an interval is given, the entire interval must satisfy the constraints. Also, if morning (AM) or evening (PM) is not specified, but the times could otherwise be regarded as correct, the response should be given the benefit of the doubt, and coded as correct.

No Credit

Code 0:   Other responses, including one time correct, but corresponding time incorrect.
- Sydney 8 am, Berlin 10 pm.

Code 9:   Missing.

Translation note: The term Internet Relay Chat (IRC) is a standard, well-known term for Internet users. Please translate this to the equivalent term used in your country, rather than translating it literally.   If you don't know what this term is, you should consult some teenagers.
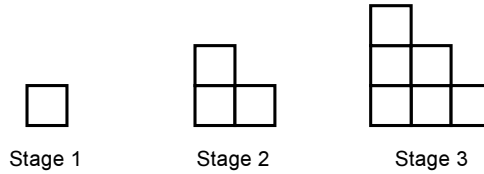
STEP PATTERN

Question 1: STEP PATTERN

Robert builds a step pattern using squares. Here are the stages he follows.

Stage 1	Stage 2	Stage 3

As you can see, he uses one square for Stage 1, three squares for Stage 2 and six for Stage 3.

How many squares should he use for the fourth stage?

Answer: .... squares.

STEP PATTERN SCORING 1

Full Credit

Code 1:   10.

No Credit

Code 0:   Other responses.

Code 9:   Missing.

List of Authors

Peter Allerup
The Danish University of Education
Tuborgvej 164, 2400 Copenhagen NV, Denmark
nimmo@dpu.dk

Dines Andersen
Danish National Institute of Social Research
Herluf Trolles Gade 11, DK-1052 Copenhagen K, Denmark
da@sfi.dk

Thomas Yung Andersen
SFI-SURVEY
Danish National Institute of Social Research
Herluf Trolles Gade 11, DK-1052 Copenhagen K, Denmark
tan@sfi.dk

Júlíus K. Björnsson
Educational Testing Institute
Borgartún 7, 105 Reykjavík, Iceland
Julkb@namsmat.is

Niels Egelund
The Danish University of Education
Tuborgvej 164, 2400 Copenhagen NV, Denmark
egelund@dpu.dk

Liv Sissel Grønmo
Department of Teacher Education and School Development
University of Oslo
Norway
l.s.gronmo@ils.uio.no

Almar M. Halldórsson
Educational Testing Institute
Borgartún 7, 105 Reykjavík, Iceland
almar@namsmat.is

Therese Nerheim Hopfenbeck
Department of Teacher Education and School Development
University of Oslo, Norway
t.n.hopfenbeck@ils.uio.no

Rita Hvistendahl
Department of Teacher Education and School Development
University of Oslo, Norway
r.e.hvistendahl@ils.uio.no

Ingmar Ingemansson
PRIM-gruppen, Stockholm Institute of Education
Box 34103, 10026 Stockholm, Sweden
ingmar.ingemansson@lhs.se

Torben Pilegaard Jensen
akf, Institute of Local Government Studies
Nyropsgade 37, 2, DK-1602 Copenhagen V, Denmark
tpj@akf.dk

Karl Göran Karlsson
Department of engineering, physics and mathematics at Mid Sweden university,
Sweden
kg.karlsson@miun.se

Marit Kjærnsli
Department of Teacher Education and School Development
University of Oslo, Norway
marit.kjarnsli@ils.uio.no

Pekka Kupari
Institute for Educational Research
P.O. Box 35, FIN-40014 University of Jyväskylä, Finland
pekka.kupari@ktl.jyu.fi

Kaisa Leino
Institute for Educational Research
P.O. Box 35, FIN-40014 University of Jyväskylä, Finland
kaisa.leino@ktl.jyu.fi

Svein Lie
Department of Teacher Education and School Development
University of Oslo, Norway
svein.lie@ils.uio.no

Lena Lindenskov
The Danish University of Education
Tuborgvej 164, 2400 Copenhagen NV, Denmark
lenali@dpu.dk

Pirjo Linnakylä
Institute for Educational Research
P.O. Box 35, FIN-40014 University of Jyväskylä, Finland
pirjo.linnakyla@ ktl.jyu.fi

Antero Malin
Institute for Educational Research
P.O. Box 35, FIN-40014 University of Jyväskylä, Finland
antero.malin@ktl.jyu.fi

Jan Mejding
The Danish University of Education
Tuborgvej 164, 2400 Copenhagen NV, Denmark
jm@dpu.dk

Ragnar F. Ólafsson
Educational Testing Institute
Borgartún 7, 105 Reykjavík, Iceland
rfo@namsmat.is

Rolf V. Olsen
Department of Teacher Education and School Development
University of Oslo, Norway
rolfvo@ils.uio.no

Astrid Pettersson
PRIM-gruppen, Stockholm Institute of Education
Box 34103, 10026 Stockholm, Sweden
astrid.pettersson@lhs.se

Beatrice Schindler Rangvid
akf, Institute of Local Government Studies
Nyropsgade 37, 2, DK-1602 Copenhagen V, Denmark
bs@akf.dk

Simon Reusch
UNI•C, The Danish IT Centre for Education and Research
Copenhagen, Denmark
sr@uni-c.dk

Astrid Roe
Department of Teacher Education and School Development
University of Oslo, Norway
astrid.roe@ils.uio.no

Karin Taube
Department of Education
Mid Sweden University, 871 88 Härnösand, Sweden
karin.taube@miun.se

Are Turmo
Department of Teacher Education and School Development
University of Oslo, Norway
are.turmo@ils.uio.no

Jukka Törnroos
Institute for Educational Research
P.O. Box 35, FIN-40014 University of Jyväskylä, Finland
jukka.tornroos@ktl.jyu.fi

Peter Weng
The Danish University of Education
Tuborgvej 164, 2400 Copenhagen NV, Denmark
weng@dpu.dk

Maria Åström
Department of engineering, physics and mathematics at Mid Sweden university
and Linköpings universitet, Sweden
maria.astrom@miun.se