Explaining reading achievement in PIRLS by age and SES

Jan Van Damme, Centre for Educational Effectiveness and Evaluation, K.U.Leuven, Belgium,
Jan.VanDamme@ped.kuleuven.be
Lobke Vanhee, Centre for Educational Effectiveness and Evaluation, K.U.Leuven, Belgium,
Heidi Pustjens, Centre for Educational Effectiveness and Evaluation, K.U.Leuven, Belgium,

Abstract

This paper investigates the effect of change in average age at the country level upon the change in average reading achievement between PIRLS 2001 and PIRLS 2006. A correction for age yields a different ranking of countries according to their average achievement. Also, the importance and the construction of an internationally acceptable SES-index will be handled.

Keywords: *PIRLS, age, socio-economic status, principal component analysis, regression-discontinuity approach*

Introduction

International studies regarding educational effectiveness give a central role to the ranking of countries as for their achievement level. Well known studies that make a comparison between countries possible are PISA, TIMSS and PIRLS. However, the average achievement score of a country depends strongly on several kinds of factors. As in other studies, PIRLS uses a multiple imputation or "plausible values (PV)" methodology to obtain proficiency scores in reading for all students, to take into account the fact that each student responded to only a part of the assessment item pool (Martin, Mullis, & Kennedy, 2007). PIRLS 2006, as PIRLS 2001, had as its target population students enrolled in the fourth year of formal schooling, counting from the first year of primary school defined by UNESCO's International Standard Classification for Education (UNESCO, 1999). Accordingly, the fourth year of formal schooling should be the fourth grade in most countries. To avoid testing very young children, however, PIRLS has a policy that the average age of children in the grade tested should not be below 9.5 years old (Martin et al., 2007). Unavoidably, in a grade-oriented sample the average age of the participants differs between countries. Concerning PIRLS 2006, this is even more pronounced as a consequence of the higher average age of two countries (Luxembourg and South-Africa). These countries participated with fifth grade students, instead of fourth grade students. Cliffordson and Gustafsson (2007) already studied the effects of age and schooling on intellectual performance. The results showed that both age and schooling increase performance. So, one of the research questions will be: do the differences between the countries as for the average age explain their differences as for their average achievement?

Another factor to take into account when countries are compared, is the equity of a country's educational system. One way to look at equity is the socio-economic status (SES) of a student, school or country and its influence on the academic achievement. SES can be measured by a large number of variables. To compare countries as for their ability to keep the gap in achievement between low SES students and high SES students small, we need a SES-index which makes a valid comparison possible. In this paper we make a suggestion for a SES-index in PIRLS.

Background and Methodology

Age and schooling effect

Table 1 shows the average age and years of formal schooling of students that participated in PIRLS 2006. As already mentioned, the average age of two countries is higher because these are fifth-grade students. But also between the other countries age differences can be observed. This was also the case in PIRLS 2001.

[Take in Table 1 about here]

The relationship between the average age and the average reading score of PIRLS 2001 and 2006 at the country level will be investigated by means of correlations. We expect that the increase or decrease in average age between 2001 and 2006 is an important explanation of the progress or decline of a country as for reading achievement between these two measurement moments.
Rindermann (2007) paid attention to the fact that international student assessment and intelligence test studies suffer from different problems of sample representativeness and statistical methods; mainly the comparability across nations is questioned. Therefore, he makes suggestions for adjustments that improve representativeness and comparability. Rindermann assumed, based on several intelligence test studies and student assessment studies, that the gain of students of a specific country is on average about 42 points in one year (refering to a scale with a mean of 500 and a standard deviation of 100, as is the case in PIRLS). Swedish studies, among others, show us that a about a third of this progress is a consequence of an extra year of age and two thirds is a consequence of an extra year of schooling (Gustafsson, 2007). This means that we have to subtract 14 points from every score for every year that students are older. The score of the two countries which participated with the fifth grade, has to be reduced with another 28 points for this extra year of schooling. We apply this suggestion to the PIRLS 2006 data to see whether this changes the ranking of countries.
The ranking of countries we obtain after such a correction is only justifiable when we can assume that aging by a year has more or less similar effects in every country. Therefore, we investigate the size of this effect of age and schooling by means of three countries who participated in PIRLS 2001 or 2006 with two different grades. For their own purposes as an additional effort, Iceland and Norway administered PIRLS 2006 to small samples of their fifth-grade students, where Sweden participated in PIRLS 2001 with grades three and four. Trying to estimate the age and schooling effect in PIRLS we looked at the regression discontinuity approach described in Luyten (2006). First we merge the files of the Swedish data from grade 3 and 4 in PIRLS 2001. As we can see in Table 2 not all students in grade 3 are born in 1991 and not all students in grade 4 are born in 1990. In further analyses we only include students with a standard school career: students in grade 3 which are born in 1991 and students in grade 4 which are born in 1990. By doing that we exclude 4.85% of the students of Sweden who participated in PIRLS 2001. This is acceptable. However, we have to take the possibility of a selection bias into account when interpreting our results.

[Take in Table 2 about here]

The following is based on the method from Luyten (2006). Grade level was re-coded to assign scores of zero to students in the lower grade and scores of one to students in the higher grade. The variable that denotes a student's date of birth is based on year and month of birth. Each date was transformed into a single number. For example, a student born in March 1990 received a score of 90.25, and a student born in April received a score of 90.33. Students born in January 1991 or later are in the lower grade, and the ones born earlier are in the higher grade. The cut-off value (91.08) was then subtracted from these scores, giving each of the oldest students in the lower grade (the comparison group) a score of zero. Table 3 illustrates the transformation of the original birth dates to the scores used in the analyses.

[Take in Table 3 about here]

Next, we explain the regression-discontinuity model expressed in equation (1). The coefficient $ß_1$ expresses the effect of age and $ß_2$ the effect of one extra year of schooling. In this equation, the effect of age is assumed linear and identical in both grades (i.e., no interaction between age and grade).

$$Y_i = ß_0 + ß_1(x_i - x_0) + ß_2 z_i + R_i \qquad (1)$$

Where:
Y_i = *reading score*
x_i = *age, pupil i*
x_0 = *cut-off value*
z_i = *grade, student i (0 if lower grade; 1 if upper grade)*
$ß_0$ = *parameter for comparison group intercept at cut-off*
$ß_1$ = *age effect*

$ß_2$ = *effect of being in the upper grade (i.e., having received an extra year of schooling)*
R_i = *random residual*

SES-index

In PISA a SES-index is constructed. It was derived from the following variables: highest occupational status of the father or mother; the highest educational level of parents converted into years of schooling; and the index of home possessions obtained by asking students whether they had at their home: a study desk, a room of their own, a quiet place to study, a computer they can use for school, an educational software, a link to the internet, their own calculator, classic literature, books of poetry, works of art, books to help with their school work, a dictionary, a dishwasher, a DVD player or VCR, the number of cellular phones, televisions, computers, cars and books at home, and three other country-specific items. In 2000 and 2003 the student scores on the index are factor scores derived from a Principal Component Analysis (PCA) which are standardised to have an OECD mean of zero and a standard deviation of one. A PCA was also performed for each participating country to determine to what extent the components of the index operate in similar ways across countries. The answer to this question turned out to be positive. In 2006 the scale construction was done through Item Response Theory (IRT) scaling with item parameters estimated first for common items separately for each country. The sum of the common items' parameters was constrained to zero for each country. Next, these item parameters were anchored and the remaining items were added, and each country was scaled separately. The correlation between the PISA 2003 and PISA 2006 indices is very high (R of 0.96). This shows that different methods of computation of the indices did not have a major impact on the results (OECD, 2007). Therefore we decide to make use of the PCA method to construct the internationally acceptable SES-index in PIRLS 2006, based on both the construction of the PISA 2003 index and the example of Vyas and Kumaranayake (2006) in constructing socio-economic status indices.

The analysis of data on socio-economic characteristics is complicated by the fact that there are potentially a large number of variables which could be measured, some of which may yield similar information. Thus a natural approach is to use methods such as principal component analysis to try and organize the data to reduce its dimensionality with as little loss of information as possible as for the total variation these variables explain (Giri, 2004).

The following variables from PIRLS 2006 can be used to construct a SES index: number of books at home[1], number of children's books at home[2], educational aids, home belongings, financially well-off[3], parents' employment situations[4], parents' highest education level[5], parents' highest occupation level[6]. The variable "educational aids" was constructed by taking the sum of the four variables: pc at home[7], own study desk[8], own books[9], daily newspaper at home[10]. These four variables were first recoded so that the answer "yes" gets a score of 1 and the answer "no" a score of 0. Using the same technique we created the variable "home belongings" based on the items 'own room'[11], 'own cell phone'[12] and four country-specific items[13]. The variables that explain too little of the variance in reading achievement and this in most countries will be left out for further analyses. For the remaining variables, the missing data will be dealt with using the expectation-maximization (EM) procedure to impute missing values. EM is a maximum likelihood procedure that works with the relationship between the unknown parameters of the data model and the missing data. This suggests an approach in which we first estimate the parameters, then estimate the missing values, then use the filled-in dataset to re-estimate the parameters, then use the

---

[1] The variable name as mentioned in the codebooks is ASBHBOOK;
[2] The variable name as mentioned in the codebooks is ASBHCHBK;
[3] The variable name as mentioned in the codebooks is ASBHWELL.
[4] The variable name as mentioned in the codebooks is ASDHPEMP;
[5] The variable name as mentioned in the codebooks is ASDHEDUP;
[6] The variable name as mentioned in the codebooks is ASDHOCCP.
[7] The variable name as mentioned in the codebooks is ASBGTA1.
[8] The variable name as mentioned in the codebooks is ASBGTA2.
[9] The variable name as mentioned in the codebooks is ASBGTA3.
[10] The variable name as mentioned in the codebooks is ASBGTA4.
[11] The variable name as mentioned in the codebooks is ASBGTA5.
[12] The variable name as mentioned in the codebooks is ASBGTA6.
[13] The variables with name ASBGTA7, ASBGTA8, ASBGTA9 and ASBGTA10.

re-estimated parameters to estimate missing values, and so on. When the process finally converges on stable estimates, the iterative process ends. EM has the advantage that it produces unbiased or nearly unbiased estimates of means, variances and covariances. For the computation of the SES-index we use the dataset with these imputated variables. Since we searched for an international measure for SES the senate weight was used to ensure that the contribution of each country is the same, regardless of the size of the population.
As we did not standardize our data, and they are therefore not expressed in the same units, we ran the analysis using the correlation matrix to ensure that all data have an equal weight. The eigen value (variance) for each principal component indicates the percentage of variation in the total data explained. The output of a PCA is a table of factor scores or weights for each variable.
Next, the average SES-score for each country was computed using the total student weight (TOTWGT) as sampling weight and the standard errors are computed using the jackknife repeated replication (JRR) method. The created variable (Y_1) can now be used as an independent variable in regression analyses with all five plausible values in turn as dependent variable. The regression coefficients and their JRR standard errors are computed (also TOTWGT was used as sampling weight).

Findings and Discussion

Age and schooling effect

The relationship between average age and average reading score at the country level is investigated by use of correlations, separately for 2001 and 2006. This correlation appears to be slightly negative (see Table 4). This means that students of well performing countries are on average slightly younger than those of less performing countries. However, the differences between countries as for average age distort the general ranking of countries as for average reading score.

[Take in Table 4 about here]

The importance of age becomes very clear when we compare the results of PIRLS 2001 and PIRLS 2006. More specific, when the relationship at the country level between the progress or decline in reading achievement between 2001 and 2006 on the one hand and the increase or decrease in average age between these two moments of testing on the other hand is considered (Table 4 and 5).

[Take in Table 5 about here]

The correlation at the country level between change in age and change in achievement amounts to 0.53. We conclude that the increase or decrease in average age is an important explanation of the progress or decline of a country concerning reading achievement. Therefore, it is recommended to make a correction for the variation in average age between countries.
The ranking of countries we obtain after such a correction (cf. Table 6) is different from the one published internationally, and from our point of view more justified, at least when we can assume that aging by a year has more or less similar effects in every country.

[Take in Table 6 about here]

The progress that students of Iceland, Norway and Sweden make in one year, equals respectively 38.65, 42.28, 41.71. On average this means that the expected progress that students make in one year can be set to 40.88. This is close to the findings of Rindermann.

As we can see in Table 7 the age effect is larger than the grade effect for Iceland and Norway, in contrast with the Swedish results as mentioned above (see Table 11). Figures 1 shows this graphically for Sweden.

[Take in Table 7 about here]

[Take in Figure 1 about here]

SES-index
In Table 8 you can find the proportion of variance of the 5 PV's explained by each of the SES variables separately and by all the variables together, and this for each country.

[Take in Table 8 about here]

In Table 9 the number of countries in which a variable explains more than a given amount of variance in reading achievement is given. Based on Table 8 and 9 we can decide for a SES index with the following variables: number of books at home, number of children's books at home, educational aids at home, parents' occupational status and parents' highest educational level. The variables that will be excluded are: family financially well off, home belongings and parents' employment situation.

[Take in Table 9 about here]

Number of children's books at home is from our point of view rather an indicator of a supportive educational environment. We also expect children's books at home to correlate highly with books at home. Therefore we suggest leaving out this variable in the construction of the SES-index. To confirm this expectation we check out the correlations between those two variables for every participating country: the correlation varies from .48 to .77 and a median equal to .64.

Based on this theoretical consideration and the statistical confirmation we can decide for a SES index with the following variables: number of books at home[14], educational aids at home[15], highest occupation level of the parents[16] and parents' highest educational level[17]. The variables that will be excluded are: number of children's books at home, family financially well off, home belongings and parents' employment situation. These four variables together explain almost the same amount of variance as the broader set of SES variables (see Table 8).

A considerable amount of countries have to deal with a lot of missing data on these variables. A few methods to handle missing data are possible. First we leave out those countries that have 50% or more missing values on at least one of the variables: United States, England and Scotland.
As already mentioned we decided to use the expectation-maximization (EM) procedure to impute missing values. We end up with a new dataset but without the three countries because of the amount of data missing (United States, England and Scotland). This dataset is used in the PCA for the computation of the SES-index.

Table 10 contains the variances explained by each component of the principal component analysis[18]. Table 11 consists of the component loadings of the first component, which are the correlations between the variables and the component. Because these are correlations, possible values range from -1 to +1.

[Take in Table 10 about here]

---

[14] Coded 1: 0.10, 2: 11-25, 3: 26-100, 4: 101-200, 5: more than 200.
[15] Ranging from 1 to 4 according to the amount of educational aids that are at the child's disposal at home.
[16] Coded 1: never worked outside home for pay, 2: general worker, 3: skilled worker, 4: clerical, 5: small business owner, 6: professional.
[17] Coded 1: some primary, lower-secondary or no schooling, 2: finished lower-secondary education, 3: finished upper-secondary education, 4: finished post-secondary but not university, 5: finished university or higher.
[18] In the PCA the four SES variables were considered as continuous variables. As for the number of 'educational aids', that is obvious. For the other three variables it is acceptable as there are five or more response categories which are ordered. The only category where this is not the case, is the one 'occupation not applicable' of the variable 'highest occupation level of the parents'; that is a rest category, which is small (<6% in 33 of the 42 countries, <11% in 40 of the 42 countries). After including this category as a separate dummy variable, the factor loading was low; for that reason the variable was excluded from the model.

[Take in Table 11 about here]

Table 12 shows the average SES-index for each country. A linear regression analysis of reading achievement can now be done on the created variable, which we call the SES-index. Note that we use the mean of the regression coefficients that were calculated for every plausible value. Table 13 shows the results. Figures 2 and 3 gives us the regression lines of several countries.

[Take in Table 12 about here]

[Take in Table 13 about here]

[Take in Figure 2 and 3 about here]

Conclusion and Implications

We tried to demonstrate the importance of age and differences between countries on the ranking of countries concerning their achievement level. We proved this statement by subtracting 14 points from every score for every year that students are older. The score of the two countries which participated with the fifth grade, has to be reduced with another 28 points. Our analyses indicated that changes in the average age between 2006 and 2001 explained an important part of the changes in average achievement of the countries between the two measurement moments.
Next, we make a suggestion for creating an internationally acceptable SES-index, by means of principal component analysis. Then, we took a look at the impact of SES on reading achievement in PIRLS 2006.

References

Cliffordson, C., & Gustafsson, J.-E. (2007). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence, 36*, 143-152..

Giri, N.C. (2004). *Multivariate statistical analysis.* New York: Marcel Dekker Inc.

Gustafsson, J.-E. (2007). Understanding causal influences on educational achievement through analysis of differences over time within countries. In T. Loveless (Ed.) *Lessons Learned: What International Assessments Tell Us about Math Achievemen*, (pp. 37-63). Washington, DC: The Brookings Institution.

Luyten, H. (2006) An Empirical Assessment of the Absolute Effect of Schooling: Regression-Discontinuity Applied to TIMSS-95. *Oxford Review of Education 32*, 397-429.

Martin, O.M., Mullis, I.V.S., Kennedy, A.M. (Eds.). (2007). PIRLS *2006 Technical Report.* Chestnut Hill: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Organisation for Economic Cooperation and Development (OECD) (2007). *PISA 2006: Science competencies for tomorrow's world. Volume 1: Analysis.* Paris: OECD.

Rindermann, H. (2007). The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS and IQ-Tests Across Nations. *European Journal of Personality, 21*, 667-706.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley

Vyas, S., & Kumaranayake, L (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy Plan, 21*, 459-468.

http://heapol.oxfordjournals.org/cgi/reprint/21/6/459

Table 1: Participating countries in PIRLS 2006 and 2001, and average age and years of formal schooling of PIRLS 2006

Countries	2006	2001	Average Age (2006)	Years of Formal Schooling (2006)
Argentina		x	~	~
Belgium Flemish	x		10.0	4
Belgium French	x		9.9	4
Belize		x	~	~
Bulgaria	x	x	10.9	4
Canada, Alberta	x		9.9	4
Canada, British Columbia	x		9.8	4
Canada, Nova Scotia	x		10.0	4
Canada, Ontario	x	x	9.8	4
Canada, Quebec	x	x	10.1	4
Chinese Taipei	x		10.1	4
Columbia		x	~	~
Cyprus		x	~	~
Denmark	x		10.9	4
Germany	x	x	10.5	4
England	x	x	10.3	4
France	x	x	10.0	4
Georgia	x		10.1	4
Greece		x	~	~
Hungary	x	x	10.7	4
Hong Kong SAR	x	x	10.0	4
Iceland	x	x	9.8	4
Indonesia	x		10.4	4
Iran	x	x	10.2	4
Israel	x	x	10.1	4
Italy	x	x	9.7	4
Kuwait	x		9.8	4
Latvia	x	x	11.0	4
Lithuania	x	x	10.7	4
Luxembourg	x		11.4	5
Macedonia	x	x	10.6	4
Morocco	x	x	10.8	4
Moldavia	x	x	10.9	4
Netherlands	x	x	10.3	4
New Zealand	x	x	10.0	4.5-5.5
Norway	x	x	9.8	4
Austria	x		10.3	4
Poland	x		9.9	4
Qatar	x		9.8	4
Romania	x	x	10.9	4
Russian Federation	x	x	10.8	4
Scotland	x	x	9.9	5
Singapore	x	x	10.4	4
Slovak Republic	x	x	10.4	4
Slovenia	x	x	9.9	3 or 4
Spain	x		9.9	4
Trinidad and Tobago	x		10.1	5
Czech Republic		x	~	~
Turkey		x	~	~
United States	x	x	10.1	4
South Africa	x		11.9	5
Sweden	x	x	10.9	4

Notes. ~ means data are not available

Table 2: Crosstabulation of grade and birth year of Swedish students in PIRLS 2001

	Grade	
Birth year	**3**	**4**
1989	12	95
1990	91	5778
1991	4987	170
1992	181	0

Table 3: Range of birth dates (cut-off = 91.08) for Swedisch students in PIRLS 2001

	Year and Month	In decimals	Minus cut-off
Upper grade	January 1990	90.08	-1.00
	February 1990	90.17	-.92
	March 1990	90.25	-.83
	April 1990	90.33	-.75
	May 1990	90.42	-.67
	June 1990	90.50	-.58
	July 1990	90.58	-.50
	August 1990	90.67	-.42
	September 1990	90.75	-.33
	October 1990	90.83	-.25
	November 1990	90.92	-.17
	December 1990	91.00	-.08
Lower grade	January 1991	91.08	.00
	February 1991	91.17	.08
	March 1991	91.25	.17
	April 1991	91.33	.25
	May 1991	91.42	.33
	June 1991	91.50	.42
	July 1991	91.58	.50
	August 1991	91.67	.58
	September 1991	91.75	.67
	October 1991	91.83	.75
	November 1991	91.92	.83
	December 1991	92.00	.92

Table 4: Correlations between average achievement, average age and change between 2006 and 2001. (*N*=28)

	Reading 2001	Reading 2006	Reading change	Age 2001	Age 2006	Age change
Reading 2001	1.00					
Reading 2006	.958**	1.00				
Reading change	.194	.468*	1.00			
Age 2001	-.051	-.040	.022	1.00		
Age 2006	-.095	-.153	-.229	.147	1.00	
Age change	.424*	.537**	.530**	-.014	.121	1.00

Table 5: Average reading achievement and average age of students for countries in PIRLS 2001 and PIRLS 2006 and change between 2006 and 2001

Countries	Achievement 2006	Achievement 2001	Difference	Age 2006	Age 2001	Difference
Russian Federation	565	528	37	10.8	10.3	0.5
Hong Kong SAR	564	528	36	10.0	10.2	-0.2
Singapore	558	528	30	10.4	10.1	0.3
Slovenia	522	502	20	9.9	9.8	0.1
Slovak Republic	531	518	13	10.4	10.3	0.1
Italy	551	541	10	9.7	9.8	-0.1
Germany	548	539	9	10.5	10.5	0.0
Moldova, Rep.of	500	492	8	10.9	10.8	0.1
Hungary	551	543	8	10.7	10.7	0.0
Iran, Islamic Rep. Of	421	414	7	10.2	10.4	-0.2
Canada, Ontario	554	548	6	9.8	9.9	-0.1
Israel	512	509	3	10.1	10	0.1
New Zealand	532	529	3	10.0	10.1	-0.1
Macedonia, Rep. Of	442	442	0	10.6	10.7	-0.1
Scotland	527	528	-1	9.9	9.8	0.1
Norway	498	499	-1	9.8	10	-0.2
Iceland	511	512	-1	9.8	9.7	0.1
United States	540	542	-2	10.1	10.2	-0.1
Bulgaria	547	550	-3	10.9	10.9	0.0
France	522	525	-3	10.0	10.1	-0.1
Latvia	541	545	-4	11.0	11	0.0
Canada, Quebec	533	537	-4	10.1	10.2	-0.1
Lithuania	537	543	-6	10.7	10.9	-0.2
Netherlands	547	554	-7	10.3	10.3	0.0
Sweden	549	561	-12	10.9	10.8	0.1
England	539	553	-14	10.3	10.2	0.1
Romania	489	512	-23	10.9	11.1	-0.2
Morocco	323	350	-27	10.8	11.2	-0.4

Table 6: Countries ranked according to reading achievement 2006 with correction for age differences

Countries	Reading Score	Age	Corrected Reading Score
Hong Kong SAR	564	10.0	568.26
Canada. Alberta	560	9.9	565.66
Canada. British Columbia	558	9.8	565.06
Canada. Ontario	555	9.8	562.06
Italy	551	9.7	559.46
Russian Federation	565	10.8	558.06
Singapore	558	10.4	556.66
Belgium Flemish	547	10.0	551.26
Netherlands	547	10.3	547.06
Canada. Nova Scotia	542	10.0	546.26
Hungary	551	10.7	545.46
Germany	548	10.5	545.26
United States	540	10.1	542.86
Sweden	549	10.9	540.66
England	539	10.3	539.06
Bulgaria	547	10.9	538.66
Austria	538	10.3	538.06
Chinese Taipei	535	10.1	537.86
Denmark	546	10.9	537.66
New Zealand	532	10.0	536.26
Canada. Quebec	533	10.1	535.86
Scotland	527	9.9	532.66
Lithuania	537	10.7	531.46
Latvia	541	11.0	531.26
Slovak Republic	531	10.4	529.66
Slovenia	522	9.9	527.66
France	522	10.0	526.26
Poland	519	9.9	524.66
Spain	513	9.9	518.66
Iceland	511	9.8	518.06
Israel	512	10.1	514.86
Luxembourg	557	11.4	513.66
Belgium French	500	9.9	505.66
Norway	498	9.8	505.06
Moldavia	500	10.9	491.66
Romania	489	10.9	480.66
Georgia	471	10.1	473.86
Trinidad and Tobago	436	10.1	438.86
Macedonia	442	10.6	437.86
Iran	421	10.2	422.46
Indonesia	405	10.4	403.66
Qatar	353	9.8	360.06
Kuwait	330	9.8	337.06
Morocco	323	10.8	316.06
South Africa	302	11.9	251.66

Table 7: Grade and age effects for Sweden, Iceland and Norway

	Grade effect	**Age effect**
Sweden	26.34***	15.54**
Iceland	10.29(*: indien eenzijdig getoetst)	28.00***
Norway	12.09	30.18***

Table 8: Linear regression of PV's on SES-variables

R^2	**books**	children's books	Well off	Home belongings	**Educational aids**	Employment situation	**Occupational Status**	**Highest Educational level**	All variables together	Selected (in bold) variables together
Austria	0.156	0.168	0.003	0.003	0.098	0.005	0.065	0.079	0.231	0.212
Belgium (Flemish)	0.070	0.084	0.012	0.015	0.061	0.008	0.090	0.139	0.197	0.187
Belgium (French)	0.119	0.151	0.008	/	0.039	0.036	0.118	0.119	/	0.200
Bulgaria	0.127	0.134	0.029	0.030	0.090	0.048	0.088	0.114	0.187	0.164
Canada, Alberta	0.033	0.046	0.018	/	0.066	0.000	0.040	0.040	/	0.109
Canada, British Colombia	0.051	0.057	0.003	/	0.066	0.000	0.040	0.038	/	0.108
Canada, Nova Scotia	0.038	0.049	0.020	/	0.068	0.009	0.050	0.075	/	0.140
Canada, Ontario	0.042	0.049	0.011	/	0.087	0.006	0.053	0.044	/	0.149
Canada, Quebec	0.057	0.060	0.022	/	0.058	0.010	0.047	0.073	/	0.123
Chinese Taipei	0.102	0.111	0.019	/	0.089	0.026	0.055	0.137	/	0.203
Denmark	0.066	0.045	0.005	/	0.028	0.008	0.064	0.064	/	0.105
England	0.089	0.085	0.039	0.002	0.082	0.011	0.061	0.120	0.179	0.166
France	0.122	0.156	/	/	0.038	0.024	0.112	0.141	/	0.197
Georgia	0.085	0.061	0.006	/	0.016	0.018	0.047	0.081	/	0.107
Germany	0.142	0.179	0.001	0.047	0.122	0.002	0.099	0.136	0.271	0.249
Hong Kong SAR	0.029	0.032	0.005	0.012	0.049	0.006	0.007	0.030	0.116	0.070
Hungary	0.186	0.199	0.023	/	0.113	0.043	0.132	0.237	/	0.287
Iceland	0.052	0.058	0.006	0.002	0.079	0.004	0.042	0.076	0.158	0.144
Indonesia	0.031	0.011	0.006	/	0.014	0.005	0.035	0.119	/	0.136
Iran, Islamic Rep. of	0.101	0.122	0.022	0.046	0.189	0.031	0.080	0.201	0.304	0.273
Israel	0.094	0.129	0.003	0.034	0.183	0.095	0.182	0.194	0.259	0.264
Italy	0.061	0.063	0.012	0.007	0.051	0.004	0.040	0.066	0.102	0.100
Kuwait	0.041	0.034	0.005	/	0.003	0.009	0.030	0.069	/	0.100
Latvia	0.066	0.102	0.009	0.024	0.060	0.008	0.062	0.052	0.164	0.135
Lithuania	0.082	0.103	0.012	0.051	0.065	0.027	0.090	0.127	0.184	0.175
Luxembourg	0.205	0.208	0.011	0.035	0.080	0.000	0.087	0.131	0.269	0.254
Macedonia, Rep. of	0.082	0.059	0.004	0.006	0.074	0.031	0.103	0.192	0.252	0.222
Moldova, Rep. of	0.059	0.045	0.008	0.011	0.033	0.011	0.021	0.055	0.105	0.092
Morocco	0.052	0.025	0.009	/	0.027	0.022	0.023	0.071	/	0.113
Netherlands	0.081	0.080	0.031	/	0.071	0.000	0.065	0.097	/	0.147
New Zealand	0.076	0.085	0.056	0.018	0.092	0.003	0.072	0.094	0.182	0.176
Norway	0.068	0.073	0.005	0.000	0.052	0.013	0.052	0.082	0.142	0.125
Poland	0.099	0.093	0.019	/	0.119	0.029	0.059	0.146	/	0.216
Qatar	0.044	0.025	0.002	/	0.001	0.019	0.036	0.076	/	0.095
Romania	0.206	0.178	0.071	0.013	0.133	0.083	0.180	0.186	0.298	0.294
Russian Federation	0.085	0.097	0.010	0.018	0.042	0.023	0.065	0.094	0.160	0.154
Scotland	0.102	0.114	0.038	0.000	0.046	0.015	0.097	0.098	0.197	0.173
Singapore	0.048	0.126	0.026	/	0.140	0.012	0.087	0.166	/	0.238
Slovak Republic	0.167	0.191	0.033	0.100	0.176	0.060	0.139	0.156	0.264	0.284
Slovenia	0.089	0.117	0.030	0.004	0.012	0.022	0.108	0.142	0.184	0.170
South Africa	0.076	0.038	/	0.176	0.127	0.110	0.102	0.193	/	0.365
Spain	0.126	0.114	0.006	0.002	0.041	0.007	0.078	0.112	0.166	0.149
Sweden	0.092	0.068	0.011	0.001	0.054	0.008	0.060	0.090	0.170	0.166
Trinidad and Tobago	0.058	0.074	0.016	/	0.065	0.019	0.077	0.124	/	0.189
United States	/	/	/	/	0.097	/	/	/	/	/

Table 9: Number of countries in which each variable explains more than a given amount in reading achievement

	books	children's books	**Well off**	**Home belongings**	Educational aids	**Employment situation**	Occupational Status	Highest Educational level
Number of countries with R^2 >.01	44	44	**23**	**13**	43	**25**	43	44
Number of countries with R^2 >.02	44	43	**12**	**7**	40	**16**	43	44
Number of countries with R^2 >.03	43	41	**6**	**6**	38	**9**	41	44
Number of countries with R^2 >.04	40	38	**1**	**3**	35	**6**	39	42
Number of countries with R^2 >.05	36	33	**1**	**1**	31	**4**	36	40

Table 10: Total variance explained by each component of the principal component analysis

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.28	56.99	56.99
2	0.75	18.87	75.86
3	0.57	14.13	89.99
4	0.40	10.01	100.00

Table 11: Component Matrix

Variable	Component loadings
Books at home	0.77
Educational Aids	0.61
Parents' highest occupational level	0.79
Parents' highest educational level	0.83

Table 12: Average SES-index (in descending order) for participating countries in PIRLS 2006

Country	SES-index
Sweden	0.74
Iceland	0.72
Norway	0.72
Denmark	0.62
Canada British Columbia	0.57
Canada Alberta	0.54
Canada Ontario	0.50
New Zealand	0.46
Canada Nova Scotia	0.46
Netherlands	0.40
Canada Quebec	0.37
Latvia	0.37
Russian Federation	0.24
Hungary	0.21
Belgium Flemish	0.20
Lithuania	0.18
Israel	0.16
Singapore	0.15
Germany	0.14
France	0.10
Slovenia	0.08
Austria	0.05
Chinese Taipei	0.05
Luxembourg	0.01
Spain	0.01
Slovakia	0.00
Georgia	-0.06
Belgium French	-0.06
Poland	-0.16
Hong Kong	-0.18
Qatar	-0.18
Italy	-0.20
Kuwait	-0.21
Bulgaria	-0.26
Trinidad and Tobago	-0.28
Macedonia	-0.53
Moldova	-0.60
Romania	-0.70
South Africa	-0.96
Iran	-1.27
Indonesia	-1.40
Morocco	-1.48

Table 13: Linear regression of reading achievement on SES-indicator, ascending according to regression coefficient

	SES-index		Reading Score**			
Country	$\bar{X}$	**SE**	$\bar{X}$	**SE**	**Regression coefficient***	**R^2**
Hong Kong	-0.18	0.04	563.91	2.35	13.61	4.8
Moldova	-0.60	0.04	499.88	3.04	22.20	9.8
Italy	-0.20	0.03	551.47	2.93	23.84	10.5
Netherlands	0.40	0.03	547.15	1.52	24.85	12.6
Qatar	-0.18	0.01	353.44	1.09	25.39	5.7
Bulgaria	-0.26	0.06	547.03	4.37	26.39	14.2
Georgia	-0.06	0.04	470.84	3.14	26.46	12.2
Spain	0.01	0.03	512.5	2.48	27.75	11.3
Lithuania	0.18	0.03	537.03	1.64	27.77	17.4
Belgium Flemish	0.20	0.03	547.04	1.96	27.93	17.5
Canada Quebec	0.37	0.03	532.91	2.76	28.02	11.4
Denmark	0.62	0.03	546.35	2.27	28.52	10.7
Canada Alberta	0.54	0.02	560.16	2.36	31.30	9.5
Norway	0.72	0.02	498.01	2.55	31.48	12.8
Latvia	0.37	0.02	540.91	2.34	31.84	13.9
Russian Federation	0.24	0.02	564.74	3.35	32.83	13.9
Chinese Taipei	0.05	0.03	535.37	2.04	32.91	19.9
Sweden	0.74	0.03	549.28	2.28	33.09	13.8
Indonesia	-1.40	0.03	404.74	4.07	33.05	10.9
France	0.10	0.03	521.29	2.07	33.25	19.3
Canada British Columbia	0.57	0.02	558.04	2.63	33.41	10.1
Iceland	0.72	0.01	510.6	1.29	33.59	11.1
Morocco	-1.48	0.04	322.58	5.94	33.42	7.5
Canada Ontario	0.50	0.03	555.19	2.73	33.59	11.1
Poland	-0.16	0.03	519.39	2.36	34.30	19.0
Luxembourg	0.01	0.01	557.2	1.08	34.32	23.2
Kuwait	-0.21	0.02	330.3	4.22	34.93	6.4
Germany	0.14	0.03	547.59	2.17	36.21	19.1
Canada Nova Scotia	0.46	0.02	541.85	2.19	36.35	12.0
Slovenia	0.08	0.02	521.53	2.09	37.53	17.2
Belgium French	-0.06	0.03	499.67	2.64	37.76	18.7
Austria	0.05	0.02	538.3	2.2	39.88	21.4
Singapore	0.15	0.02	558.27	2.88	40.22	21.3
Hungary	0.21	0.04	550.89	2.98	40.53	27.9
Slovakia	0.00	0.03	530.81	2.75	42.68	26.0
Iran	-1.27	0.05	420.93	3.09	44.40	24.6
Romania	-0.70	0.05	489.47	5.01	46.32	27.2
New Zealand	0.46	0.02	531.72	2.02	47.15	14.4
Trinidad and Tobago	-0.28	0.03	435.59	4.89	49.28	15.9
Macedonia	-0.53	0.04	442.4	4.09	50.14	20.4
Israel	0.16	0.02	512.46	3.35	52.78	19.8
South Africa	-0.96	0.03	301.61	5.56	67.77	21.1

* The mean of the regression coefficients of all five plausible values.

** The mean of the means of all plausible values. Standard deviations are calculated according to the following formula: $SD = \sqrt{T}$ where:

- T = within-variance + 1.2 * between-variance
- Within-variance = (mean of *m* standard errors)²
- Between variance = 0.25* S(one plausible value – mean of *m* plausible values)²
- *m* = number of residuals

(Rubin, 1987).

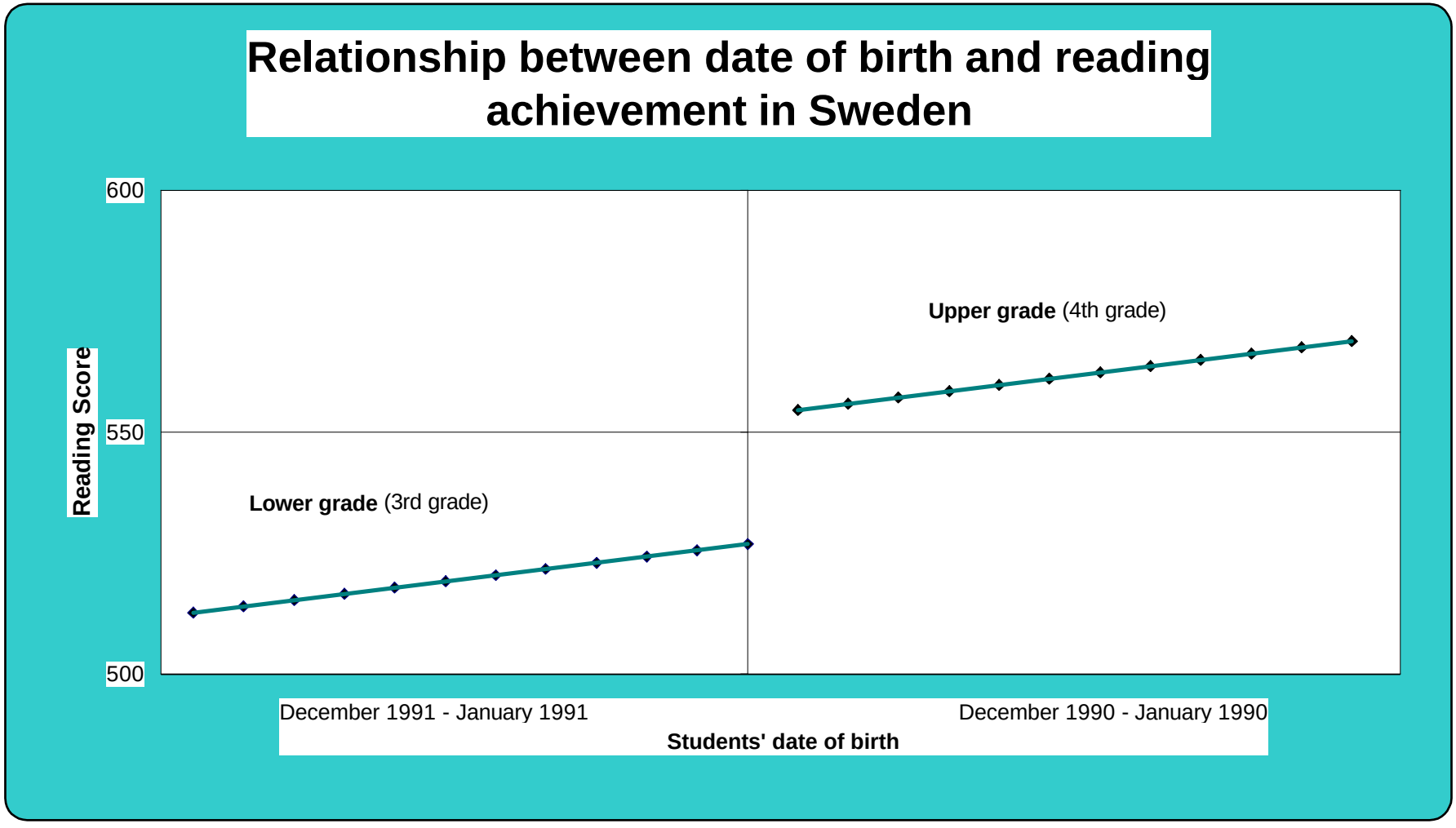

Figure 1. Relationship between date of birth and reading achievement in Sweden.

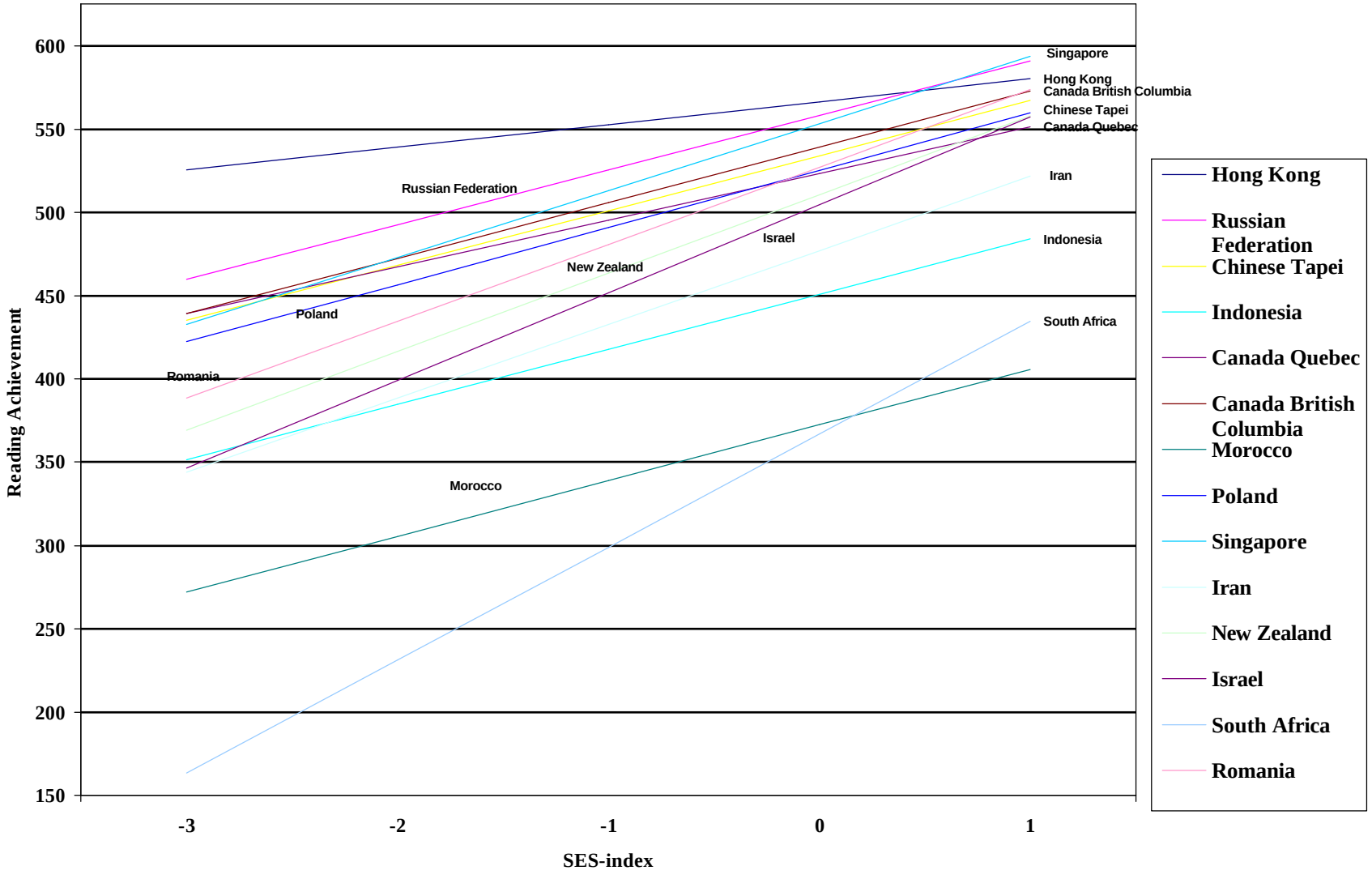

Figure 2: Linear regression of reading achievement on SES.

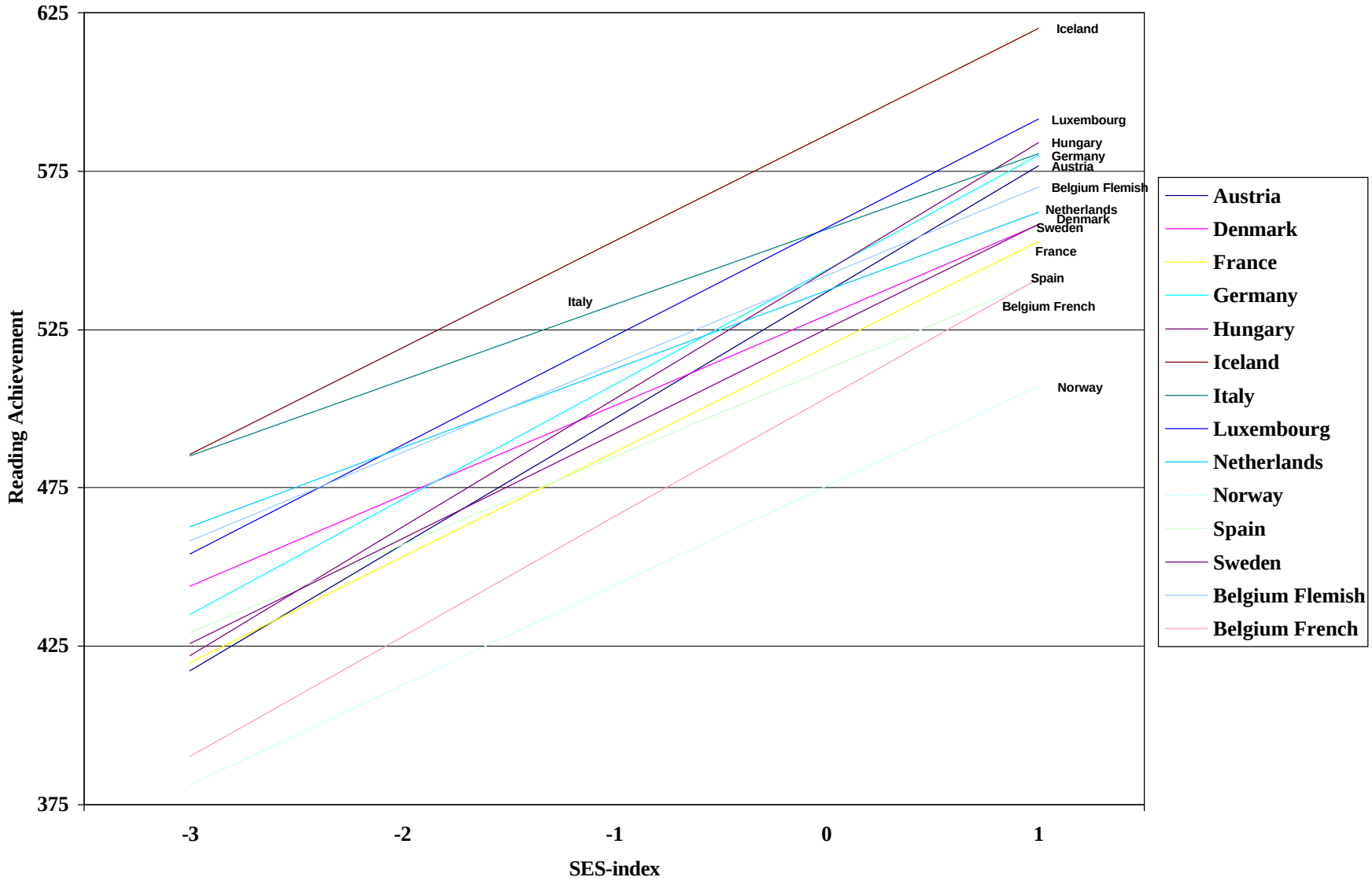

Figure 3: Linear regression of reading achievement on SES, West-European countries.