Edukacja 2017, 2(141), 100–112
doi: 10.24131/3724.170206
issn: 0239-6858

Differentiation of initial skills in 1st grade and the relative age effect in 3rd grade

Krzysztof Konarzewski
Educational Research Institute*

The paper aims to verify two hypotheses. The first states that the differentiation of initial 1st grade pupil skills is negatively correlated with age at entry to school. This was tested according to findings from the ERI study Six and seven year olds at the start of school. The variation of reading and writing scores of 1164 pupils aged 5.9–7.9 years who were beginning 1st grade decreased in subsequent age quartiles. The second hypothesis – that classroom differentiation of pupils' initial skills in 1st grade is positively correlated with the birthdate effect in 3rd grade – was tested on data of 4838 pupils from 254 classrooms drawn from the 2011 PIRLS and TIMSS studies in Poland. Skills were evaluated on the basis of parental reports. Hierarchical linear analysis (gender, SES and school location controlled) showed that the greater the differentiation of language skills in 1st grade, the greater the birthdate effect in mathematics in 3rd grade. This result suggests that school entry age is of lesser importance than the methods used to reduce differences in children's school readiness at the onset of education.

Keywords: education; school readiness; birthdate effect; PIRLS; TIMSS; HLM.

The aim of this article is to explain an intriguing phenomenon: a small but statistically significant part of scholastic achievement of ten-year-old pupils from 25 European countries is dependent on the interaction of the relative and absolute age of pupils (Konarzewski, 2013). The interaction is illustrated in Figure 1. Individual points represent average mathematics achievement of children divided into groups by three criteria: average age at school entry (6.2, 6.6, and 7.1 years), relative age in the class (younger, middle-aged, older) and gender. We can see that the achievement of the oldest pupils in their classes is higher than the youngest pupils, and this difference is the greatest in early-start classes, while in late-start classes, it is indistinguishable from zero. The differences in achievement of girls and boys tend to diminish in a similar way. The average age of pupils in the class depends on legal regulations and social standards. For this reason, the interaction can be detected only in international data.

Any explanation of this interaction must be embedded in the time perspective. Many studies show that the difference between achievement of older and younger pupils in a class (referred to as the *relative age effect*) decreases in consecutive years of education (DiPasquale, Moule and Flewelling, 1980; Dolata and Pokropek, 2012; Hutchison and Sharp, 1999, after: Sharp, 2002; Jones and Mandeville, 1990; Langer, Kalk and Searls, 1984; Verachtert, De

* Address: ul. Górczewska 8, 01-180 Warszawa, Poland.
E-mail: konarzewski@neostrada.pl

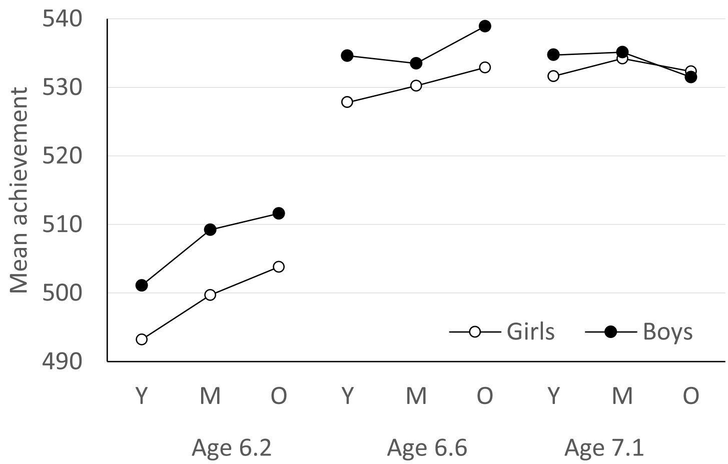

Figure 1. Mean achievement in mathematics for girls and boys of various relative age (Y – younger, M – middle-aged, O – older) in classes of different school starting age.

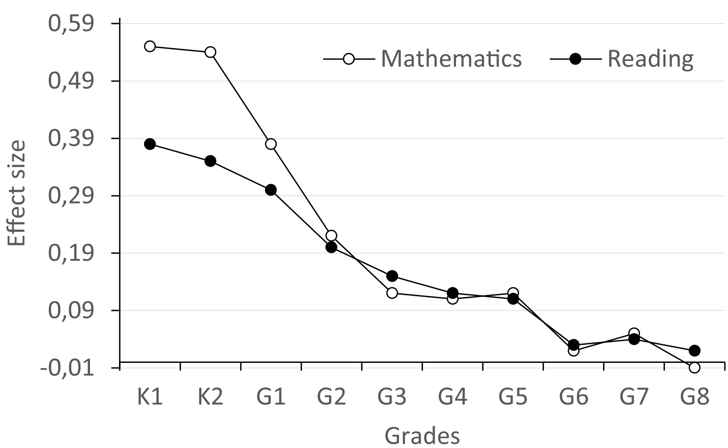

Figure 2. Differences between achievement of older and younger pupils in kindergarten (two repeated measurements) and in grades 1–8.

Based on: Oshima and Domaleski (2006).

Fraine, Onghena and Ghesquière, 2010). Figure 2, taken from the lucid research of T. Chris Oshima and Christopher Domaleski (2006), shows the differences[1] in achievement of older pupils (born within a three-month period before the cut-off for school entry) and of younger pupils (born within a three-month period after the cut-off). The data for grade K comes from longitudinal measurements in the autumn and spring of the same year, and data for grades

[1] In $d = (m_s - m_m)/s$ units, where ms and mm are the average achievement of older and younger pupils and s is the pooled standard deviation.

1–8 are cross-sectional. Each point in the figure represents the result of the comparison of two samples, approximately 3000 pupils each. As we can see, at the beginning of education, the differences are considerable (particularly for mathematics), but they tend to decrease and from 6th grade on, they are statistically insignificant.

Why does the difference between the achievement of younger and older children in the class tend to decrease? The answer seems straightforward: evidently, school creates favourable conditions for the faster development of younger rather than older children. And why does this difference tend to decrease at a rate that depends on school starting age? The consistent answer is: because the younger the children are at the start, the harder it becomes to create such conditions. What does this difficulty consist of? The hypothetical answer consists of two parts. First, the lower the school starting age of 1st grade children, the larger the differences between them in terms of initial skills. Second, the larger the differences, the harder it is for the teacher to adapt the modal level of developmental challenges to the needs of those less advanced (including younger children) in a way that does not affect those who are more advanced (including older children). Because of this difficulty, younger children tend to develop at a rate that is slower than usual and catching up with older children takes them more time. This is why the age levelling rate depends on school starting age. We may use the statement provided above to draw two hypotheses:

- H_1: The differentiation of initial skills in an average 1st grade class is negatively correlated with school starting age.
- H_2: The differentiation of initial skills in an average 1st grade class is positively correlated with the relative age effect in 3rd grade.

This article aims to verify both hypotheses.

School starting age and the differentiation of initial skills

Numerous studies prove the existence of mental dispositions whose level at school entry predicts (to a statistically significant but not necessarily practically important degree) future school achievement[2]. Typical research generally estimates the coefficient of regression of achievements in a schooling period to dispositions measured in a pre-school period such as: (a) initial language and numeracy skills, (b) cognitive dispositions, e.g. overall intelligence, capacity of short-term and working memory, executive functions, (c) affective and motivational dispositions, e.g. resilience. The dispositions are measured with the help of psychological tests and scales developed either by the research author (e.g. Wilgocka-Okoń, 2003) or by other researchers. The Phelps Kindergarten Readiness Scale (Augustyniak, Cook-Cottone and Calabrese, 2004) or the Woodcock-Johnson Psychoeducational Battery (Evans, Floyd, McGrew and Leforgee, 2002) may serve as examples here. There may also be a set of separate tests (e.g. Bull, Espy and Wiebe, 2008; Gathercole, Alloway, Willis and Adam, 2006) or sub-tests selected from various tools (Konold and Pianta, 2005; Kurdek and Sinclair, 2001) and even scholastic achievement tests (Princiotta, Flanagan and Hausken, 2006).

The large number of such studies induces the production of meta-analyses. Karen La Paro and Robert Pianta (2000) performed one on 70 older reports. The frequently cited article of Greg Duncan and associates (2007) presents meta-analysis results for six American, Canadian and British large-scale studies. The set of independent variables consisted of the measurements of language

[2] These studies are usually categorised as school readiness – a highly complex and problematic term, which is not needed in this article. The role of school readiness in the theory of education and educational practice deserves separate critical treatment.

and numeracy skill, attention, emotions and social behaviour at the start of school (children aged 5–6). In five of these studies, the data of children aged 3.5–4 were collected, but they were used only as covariates in the meta-analyses. Dependent variables consisted of school achievement in reading and mathematics, teachers' assessments of pupils' characteristics and grade retention data of children aged 8–14. None of these studies measured the effects of learning in 1st grade. The strongest predictors of future school achievement in mathematics and reading turned out to be initial numeracy skills (the weighted mean regression coefficient was 0.33). Ranked in second place were initial language skills (0.13), which were more prognostic of the achievement in reading than in mathematics. Third place was taken by attentional measures (0.07). Measures of initial behavioural and social skill-related problems did not differentiate future school achievement. The results did not depend either on gender or on the socioeconomic status (SES) of pupils' families. They did not depend on whether the pupil data source consisted of standardized tests or teachers' opinions.

The studies fail to provide any information on the differentiation of initial dispositions in age groups. What could one expect in this case? Let's assume that we know a certain number of dispositions predicting future school achievement and that we can define a threshold value on each scale of these dispositions – so that the lower (and sometimes higher) values can be recognised as risk factors for failure in learning (i.e. probabilistic inversions of preconditions for successful learning). Brzezińska et al. (2014) refer to them as deficiencies and surpluses. The initial dispositions of each pupil can be presented as a profile. Figure 3 shows a pupil's profile for four variables: A–D. The simplest (though not necessarily the most valid) measurement of a pupil's risk of failure is the number of measurements (in this case: 1) in which he or she has ranked below the threshold value. Assuming that children, either under the influence of maturation or experiences (including those invoked by their teacher), achieve increasingly higher levels of variables A–D at an individual rate and that these progressions are relatively irrevocable, the total curve of initial development in a population should monotonically increase with age and its dispersion should reach its maximum in the middle of this process. This means that a country shifting

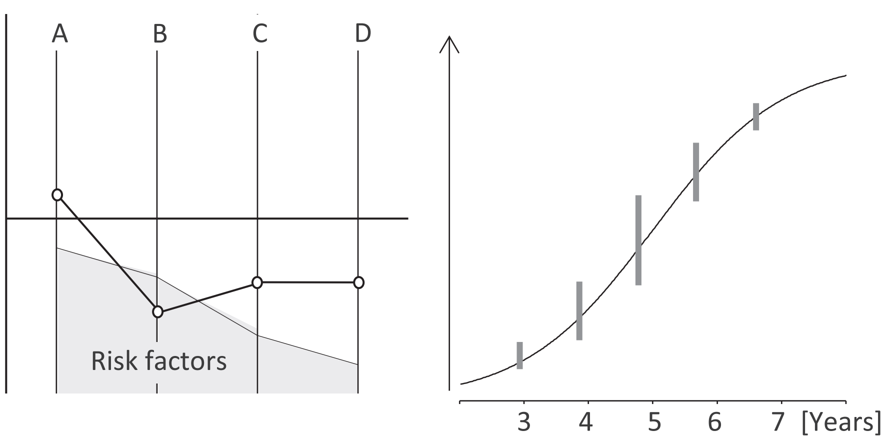

Figure 3. A profile of initial skills A–D (left) and a total developmental curve (right).

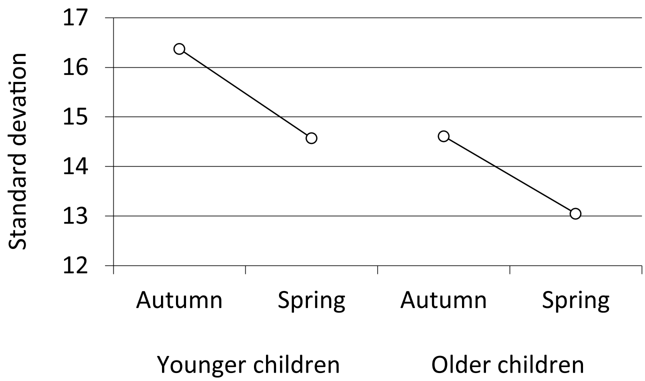

Figure 4. Standard deviations of younger and older children's reading skills in kindergarten.

Based on: Oshima and Domalewski (2006).

the school starting age from 7 to 6 should be aware of the increased differentiation of initial skills in a cohort of first-graders.

The way of thinking presented above leads to hypothesis H_1, which states that the differentiation of initial skills is negatively correlated with the age of school entry. The survey of the literature has provided only three reports that specify initial skill dispersion across various age groups of children. The team of Laura DeThorne (2010) presented the results of double measurements (at an annual interval) of six elementary aspects of reading in a group of 380 six-year-olds. With the exception of the Stanford-Binet word definition sub-test, all dispersions in the second measurement, despite higher means, were lower than those in the first measurement. Christophe Mussolin and others (2014) conducted two tests (at a 7-month interval) of the use of numerical symbols and quantitative comparisons by a group of children aged 3–4. Four in five measures had lower dispersions (two of them significantly lower) in the second measurement. Higher dispersion was only observed in the "count to 60" task, which was quite difficult for the children as they managed to reach (on average) 13 in the first and 21 in the second measurement. Oshima and Domaleski (2006) report dispersions of younger and older children's reading and mathematics skills in a preparatory class. Figure 4 shows dispersions of the reading measurements. The decrease of the dispersions within and between the two groups is statistically significant ($p < 0.001$). For mathematics, a significant reduction of dispersions within the groups was observed, but there were no differences between the groups.

In order to verify hypothesis H_1, data taken from *Badanie sześcio i siedmiolatków na starcie szkolnym* [*Six-year-olds and seven-year-olds at school entry level*] study carried out by Radosław Kaczan and Piotr Rycielski (2014) were used. This study aimed to define the growth of skills of six- and seven-year-olds within the seven months spent in one of four milieus: a preparatory class in kindergarten, a preparatory class in a school, in a 1st grade class, and in a 2nd grade class. The study subjects were 3029 pupils selected on the basis of their personal identification (PESEL) number in proportion to the size of the voivodships. An adaptive Test of Skills at the Start of School (*Test umiejętności na starcie szkolnym*, TUNSS) was used to measure initial skills in reading, writing and mathematics. The mean of the results was

Table 1
Means and standard deviations of the measurements at the beginning of 1st grade (I) and towards the end of 1st grade (II) for age quartiles

Measurements		Quartiles of age 1 (6.1 years; N = 133)	2 (6.6 years; N = 242)	3 (7.0 years; N = 445)	4 (7.6 years; N = 292)	Mean/ /variance tests
Mathematics						
I	Mean	93.9	97.0	99.0	100.9	<0.001
	Standard deviation	12.0	13.0	12.2	11.3	ns.
II	Mean	104.7	105.9	105.9	107.4	<0.001
	Standard deviation	11.4	11.4	11.3	11.2	ns.
Reading						
I	Mean	93.7	96.6	97.9	99.9	<0.001
	Standard deviation	13.3^{*}	12.3^{*}	11.3^{**}	10.9^{**}	<0.001
II	Mean	109.5	109.9	108.7	110.9	0.010
	Standard deviation	9.6	9.2	9.4	8.2	ns.
Writing						
I	Mean	99.6	100.8	101.1	104.2	<0.001
	Standard deviation	11.8^{*}	12.1^{*}	11.5^{*}	10.4^{**}	0.001
II	Mean	108.3	111.5	110.1	113.6	<0.001
	Standard deviation	12.0^{*}	12.2^{*}	10.7^{**}	8.6^{***}	<0.001

Standard deviations with different superscripts differ significantly from each other. Own calculations based on: Kaczan and Rycielski (2014).

set to 100 and the standard deviation to 15. During the first measurement carried out in the autumn of 2012, the age of children ranged from 5.9 to 7.9. They were divided into quartiles of approximately six months of life. Data on 1164 children starting school in 1st grade were used to verify the hypothesis. The second measurement included 1112 of these children.

Means and standard deviations of the measurements at the beginning of 1st grade and towards the end of 1st grade for age quartiles are shown in Table 1. The analysis of variance of the repeated measurements shows that: (a) older children achieved better results than younger children ($p < 0.001$); (b) in the second measurement, an increase in the results was observed ($p < 0.001$); (c) the size of the increase in mathematics and reading was dependent on age: younger children made better progress than older children ($p < 0.001$). A comparison of the variances using Levene's test indicates that: (a) variances of both measurements of mathematics skills for the age quartiles are equal; (b) in the first measurement of reading skills, the variances for the group of younger children (born in 2006) are significantly higher than those for the group of older children (born in 2005); in the second measurement, however, they are equal; (c) in the first measurement of writing skills, the variance drops only for the group of the oldest children; in the second measurement, however, it decreases for both quartiles of 2005.

Figure 5 helps capture the meaning of the observed changes. In the first measurement of reading skills, there is a clear age effect both in the means and standard deviations.

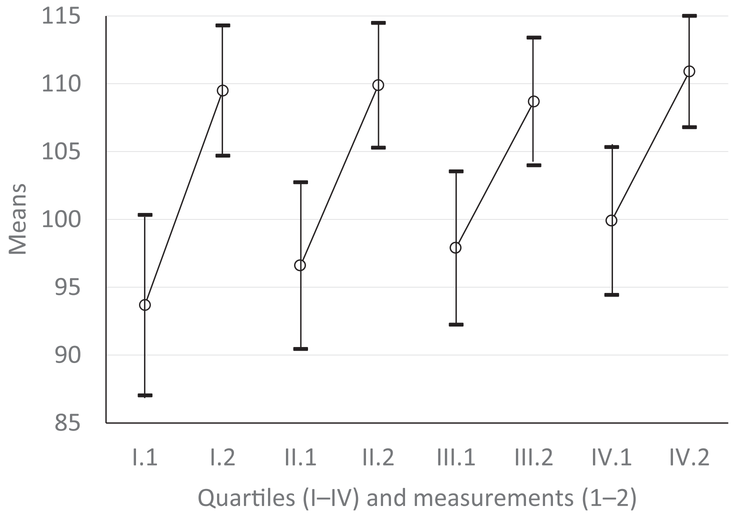

Figure 5. Means and standard deviations of two measurements of reading skills of four quartiles of age.

Own calculations based on: Kaczan and Rycielski (2014).

However, it took just seven months for the age effect to fade away from the means and to totally disappear from the standard deviations, and for the standard deviations to significantly decrease compared to the first measurement. Changes such as decreased differentiation of children's initial skills should be expected from a good education.

Differentiation of initial skills and the relative age effect in 3rd grade

In order to verify hypothesis H_2, data were used from the 2011 PIRLS (Progress in International Reading Literacy Study; Mullis, Martin, Foy and Drucker, 2012) and 2011 TIMSS (Trends in International Mathematics and Science Study; Mullis, Martin, Foy and Arora, 2012) studies carried out in Poland (Konarzewski 2013). 111 records of pupils born in the years other than 2001, 60 incomplete records, one class comprising less than five pupils and two classes (one of five and one of six pupils) with extreme regression coefficients of age effect on initial skills and with extreme residuals were removed. Following these reductions, the analysis included the data of 4838 pupils

Table 2
Structure of the sample

	Classes		Pupils	
School location	*N*	%	*N*	%
Village	79	31	1 280	26
Small town (up to 20 000 residents)	48	19	973	20
Average town (20 000–100 000 residents)	57	22	1 181	24
Large town (over 100 000 residents)	70	28	1 404	30

Table 3
Parental response distributions (%)

Could your child do the following when he/she began primary school?	*N*	Very well	Average	Poorly	Not at all
Recognise most of the letters of the alphabet	4 900	47	39	13	2
Read some words	4 863	34	39	20	7
Read sentences	4 838	14	36	30	21
Write letters of the alphabet	4 867	33	45	20	3
Write some words	4 860	24	41	25	10
	N	To 100 or more	To 20	To 10	Not at all
Count by himself/herself	4 911	30	44	25	1
	N	More than 4	3–4 figures	1–2 figures	None
Recognise different shapes (e.g. square, triangle, circle)	4 909	45	41	13	1
	N	All	5 –9 numbers	1–4 numbers	None
Recognise written numbers from 1–10	4 885	79	11	8	2
Write numbers from 1–10*	4 621	71	13	12	4
	N	Yes	No		
Do simple addition	4 902	84	16		
Do simple subtraction	4 818	71	29		

* 217 missing data records were replaced with the mean.

(96%) from 254 3rd grade classes (99%). Table 2 shows the sample structure.

Dependent variables were the achievements in mathematics and reading represented by five plausible values while independent variables were the class aggregates (means and variances) of initial skills assessed on the basis of parental response. Parents were asked to define the language and numeracy skills of their children at school entry (Table 3). In view of the high internal consistency of the responses (Cronbach's $\alpha = 0.92$ for language skills and 0.80 for numeracy skills) and the high level of unidimensionality (73% and 52% of common variance in the first component of PCA), two scales were created with the help of the single-parameter IRT model. On the individual level, the distributions of both scales are skewed to the left and are correlated with each other ($r = 0.61$). In the analysis, they are entered only as class aggregates: means and variances.

The initial skill variance in 1st grade includes the age effect ($d = 0.34$ for language skills and 0.25 for numeracy skills) and for this reason, it cannot play the role of the age effect predictor in 3rd grade. To eliminate the age effect from the initial skill variance, a two-level regression model was estimated: $Skill_{ij} = (\gamma_{00} + u_{0j}) + (\gamma_{10} + u_{1j})\, Age_{ij} + r_{ij}$. The residual r_{ij} is a new measure of $Skill_{ij}$. Pupils' residuals in classes no longer include the covariance of age and skill[3]. Classroom means for both skill measures are positively correlated ($r = 0.69$), as are the residual variances (0.44).

[3] Gary King (1986) strongly criticised the regression of residuals. However, his criticism relates to a specific case: replacing the equation $Y = \beta_1 X_1 + \beta_2 X_2 + e$ with the equation $e_1 = \beta'_2 X_2 + e_2$, where e_1 is a residual from equation $Y = \beta'_1 X_1 + e_1$. Estimates of β'_1 and β'_2 are then biased due to the omitted variable (X_2) in the last equation. Note that the analysis performed here is different: it removes the controlled variable effect from the independent variable. Omitted-variable bias is not so certain any longer.

The correlations between means and variances are negative (from -0.31 to -0.41).

The analysis was performed with the help of a two-level hierarchical linear regression (Raudenbush and Bryk, 2002). On the first level – within each of 254 3rd grade classes, the following coefficients were estimated:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\, Age_{ij} + r_{ij}, \qquad (1)$$

where Y_{ij} is the result of the reading or math test of pupil *i* from class *j*; *Age* – class centred pupil's age, i.e. a pupil's age (in years) less the average age in the pupil's class; β_{0j} – class constant (due to centring, it is equal to the average of achievement *Y* in a given class), β_{1j} – class slope, i.e. the relative age effect in class *j*, and r_{ij} is the individual residual, i.e. the difference between the expected and observed results of pupil *i* in class *j*.

On the second level (between-class), the following coefficients were estimated:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, m(Language)_j + \gamma_{02}\, m(Number)_j + u_{0j}, \qquad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, var(Language)_j + \gamma_{12}\, var(Number)_j + u_{1j}. \qquad (3)$$

The first equation depicts the dependency of the classroom mean of achievement on the classroom mean of initial language skills *m*(*Language*) and numeracy skills *m*(*Number*). Coefficients γ_{01} and γ_{02} are the measures of the dependency. Coefficient γ_{00} is an estimated value of average achievement in the population (*grand mean*) and u_{0j} is a random residual. The second equation is designed to test the dependency of the age effect in each class on the variance of initial language skills *var*(*Language*) and numeracy skills *var*(*Number*) in the given class. Coefficient γ_{11} and γ_{12} are the measures of the dependency. The statistical significance of these coefficients will confirm hypothesis H_2. Coefficient γ_{10} is an estimate of the average age effect in the population and u_{1j} is a class residual for this equation.

The equations listed above also included controlled variables. On level 1, it was gender and SES of a pupil's family (cf. Konarzewski, 2012, p. 64), on level 2 – the location of the school (with the help of three dummy variables which express the contrasts between a small, average and large town and village) and mean classroom SES. The overall model containing only significant predictors with adapted subscripts can be found in Table 4. Predictors – except for age and the dummy variables – were standardized, as their original scales do not transmit any information. This is not synonymous with the standardization of regression coefficients.

Results

The results of the analysis are provided in Table 4. We will begin our comments with the controlled variables. As shown, both gender and SES of a pupil's family significantly differentiate achievement. On average, boys are better than girls by 8.8 points in mathematics (d = 0.14) and girls are ahead of boys by 14.7 points (0.24) in reading. A difference in SES by one standard deviation makes a difference of 23.6 points (0.38) in math and 24.9 points (0.40) in reading. Class means of achievement are associated with the location of the school: rural schools are ahead of municipal schools, particularly in reading. This result ceases to surprise us if we take into account that it appears only when the SES of pupils' families is controlled. Without considering this variable, schools in small municipalities still have worse reading results compared to villages, but in large cities, the results are considerably better both in reading and mathematics. This means that the advantage of urban schools, as shown in the annual reports of the Central Examination Board, derives merely from the fact that the size of a municipality is correlated with the average SES of pupils' families. Note that in the analysis, SES is represented by two variables: individual and mean per class. They have different meanings and play different roles in the model. The first one provides

Table 4
Results from the two level regression model

	Mathematics			Reading		
Fixed effects	Coefficient	*se*	*p*	Coefficient	*se*	*p*
Mean achievement [γ_{00}]	485.96	2.85		544.72	3.00	
School location						
Small town–village [γ_{03}]	-5.46	3.87	ns.	-13.50	4.12	0.001
Average town–village [γ_{04}]	-9.61	4.17	0.022	-9.86	4.16	0.018
Large town–village [γ_{05}]	-7.50	4.03	0.064	-11.77	4.26	0.006
Mean of initial numeracy skills [γ_{02}]	6.33	1.55	<0.001	5.19	1.58	0.001
Mean of SES [γ_{01}]	22.50	1.99	<0.001	23.09	1.70	<0.001
Gender [γ_{10}]	8.79	1.94	<0.001	-14.70	1.88	<0.001
SES [γ_{20}]	23.59	1.26	<0.001	24.94	1.27	<0.001
Age [γ_{30}]	21.29	3.60	<0.001	20.29	3.61	<0.001
Variance of initial language skills [γ_{31}]	8.54	4.20	0.047	6.37	3.67	0.082
Random effects						
Variance of mean achievement [τ_{00}]	192.78		<0.001	208.13		<0.001
Variance of residuals [σ^2]	3 873.02			3881.97		
Percent of variance σ^2 reduction	10			12		
Percent of variance τ_{00} reduction	74			72		

Model: $Y_{ij} = \gamma_{00} + \gamma_{01}\, m(SES)_j + \gamma_{02}\, m(Number)_j + \gamma_{03}\, S_j + \gamma_{04}\, A_j + \gamma_{05}\, L_j + \gamma_{10}\, Sex_{ij} + \gamma_{20}\, SES_{ij} + \gamma_{30}\, Age_{ij} + \gamma_{31}\, var(Language)_j\, Age_{ij} + u_{0j} + r_{ij}$. Standard errors are robust.

information on the material and non-material resources of a pupil's household and accounts for 9–10% of the variance of individual residuals. The second indirectly provides information on the resources of the community in the region of the school and – consequently – on the equipment and normative climate of the school. When added to the model, it reduces the variance of mean achievement by 63–64%. Such decomposition of a single variable is a strength of the applied method of analysis.

Let us consider the age effect. The coefficients γ_{30} provide information on this. When gender, SES and school location are controlled, a one-year difference adds 21.3 points (d = 0.32) to the mathematics achievement of older children and 20.3 points (0.33) to their reading achievement. In educational studies, these differences are considered moderate and they are definitely not to be ignored. However, the level of confidence in predicting these differences is low – for a specific class, they may have a higher, lower or negative value. In this sense, blaming the education system for condemning younger children to failure is not justified. Please keep also in mind that this analysis is limited to normative age children, i.e. appropriate for their class. If we had taken into account all children, as in Figure 6, the age effect would be smaller. In the figure, the dark points representing children born in subsequent quarters of 2001 form a clear, growing trend. The two points on the left represent children born in later quarters who attended school one year earlier. Their achievement is higher than those of the youngest pupils at their normative age. The four points on the right represent children born in earlier quarters, i.e. the oldest ones. They were attending 3rd grade either because

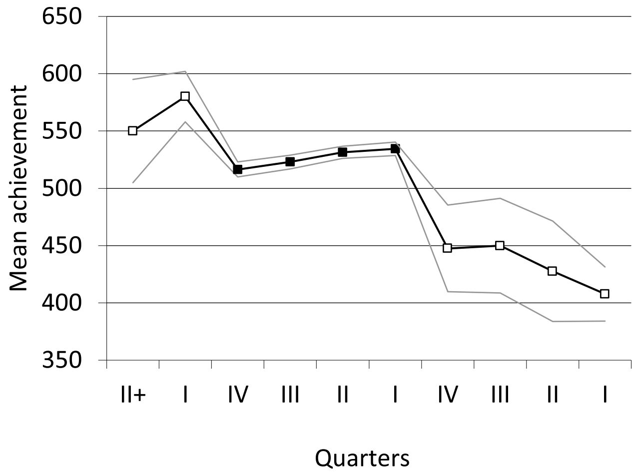

Figure 6. Weighted means of achievement by age of pupils (in quarters of the year of birth). 95-percent confidence interval set out with the jackknife method.

their school entry had been postponed or they had repeated a grade. Their achievement is much lower than those of the oldest children at their normative age. After including the full data set in the analysis, the relative age effect in reading decreases from 0.33 to an insignificant value of 0.07. While the role of age in education should be acknowledged, it should not be overestimated.

Finally, let us revise the most crucial part of the analysis showing the initial skill effect. First, the numeracy skill mean in a class predicts the class mean of achievement. After three years, the classes where children (according to their parents) were better at mastering arithmetic rudiments have substantially higher achievement in both mathematics and reading. The language skill mean turned out to have no significance in terms of future achievement.

Second, the initial language skill variance in a class predicts the age effect per class. For highly differentiated classes in terms of initial reading and writing skills, the age effect in 3rd grade is higher than in classes with lower differentiation. Coefficient γ_{31} indicates that the increase of the initial differentiation of language skills by one standard deviation significantly increases the age effect in mathematics towards the end of 3rd grade by 8.54 points. In terms of reading, this increase is 6.37 points and is found just under the threshold of significance. Language skill differentiation accounts for 13% of age difference variance in mathematics and much less (only 4%) in reading. Numeracy skill differentiation turned out to have no significance for future achievement.

Discussion

The presented results do not provide sufficient confirmation of the hypothetical explanation of school achievement dependency on the interaction of relative and absolute pupil age. Hypothesis H_1 – featuring the negative correlation of initial skill differentiation with age – is supported by the results of few published studies. To a certain extent, this may result from a negligence in reporting information on the dispersion of the measurements. The Kaczan and Rycielski (2014) study confirms the hypothesis, but only with reference to reading and writing skills. The answer to the question, "Why do the dispersions of mathematics test results fail to decrease?" remains unknown. The mathematics skill development curve (see Figure 3) may differ from other development curves primarily by a lower slope and delayed plateau; in standard language, mathematical development may be referred to as a "long march". Only an independent study may prove whether or not this is really so.

The second hypothesis referring to the relation between initial skill differentiation and the relative age effect in 3rd grade has been confirmed in a doubly limited scope. The first limitation results from basing a child's skill assessment on parents' memories which – although reliable – surely provide limited validity. It is decreased by an inclination towards presenting one's own child in a good light and towards adapting these memories to the child's current position in school. The former is impossible to estimate, but the latter seems not too large, as the inclusion of initial skill measures in the first level equation reduces the variance of achievement residuals r_{ij} only by 7.5%.

The second limitation of the scope of the hypothesis is due to the fact that age effect is significantly associated only with the differentiation of initial language skills. Why has a similar association not been observed in terms of numeracy skill differentiation? This question is related to another question: Why is 3rd grade achievement predicted only by the mean of initial numeracy skills? The hypothetical answer to both questions follows. The initial language skill variable is the indirect and approximate measure of a child's communication competency and the initial numeracy skill variable – an approximate measure of a child's intelligence. The higher the average intelligence of children in a class, the higher will be their average achievements in both mathematics and reading, regardless of the teaching methods. However, a differentiation of communication competences implies a differentiation of teacher–pupil interactions. The teacher may involuntarily prefer interactions with more attractive partners, for example, engaging with them in extended oral exchanges (see Konarzewski, 1993) to the detriment of those who are less attractive. This is equivalent to shifting the modal level of developmental challenges towards more advanced pupils (including older ones), thus decelerating the younger children's process of catching up with older children. The reader may ask, "Why should the differentiation of intelligence not work in a similar way?". The answer to this question may refer to a professional standard that prohibits teachers from devoting more of their time to more able children than to less able ones. Teachers who are aware of this standard may have better control over their inclination towards investing their efforts where higher profits are more easily gained.

The present study leaves many issues open. It confirms, however, the conclusion of earlier analyses (Konarzewski, 2012): the question about the age in which children should start school is of less importance than the question, "In what way can within-class differences in initial skills be reduced?".

Literature

Augustyniak, K. M., Cook-Cottone, C. P. and Calabrese, N. (2004). The predictive validity of the Phelps Kindergarten Readiness Scale. *Psychology in the Schools*, *41*(5), 509–516.

Brzezińska, A. I., Jabłoński, S. and Ziółkowska, B. (2014). Specyficzne i specjalne potrzeby edukacyjne. *Edukacja*, *127*(2), 37–52.

Bull, R., Espy, A. K. and Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228.

DeThorne, L. S., Petrill, S. A., Schatschneider, Ch. and Cutting, L. (2010). Conversational language use as a predictor of early reading development: language history as a moderating variable. *Journal of Speech, Language & Hearing Research*, *53*(1), 209–223.

DiPasquale, G. W., Moule, A. D. and Flewelling, R. W. (1980). The birthdate effect. *Journal of Learning Disabilities*, 13, 234–238.

Dolata, R. and Pokropek, A. (2012). Czy warto urodzić się w styczniu? Wiek biologiczny a wyniki egzaminacyjne. In B. Niemierko and M. K. Szmigel (eds.), *Regionalne i lokalne diagnozy edukacyjne* (pp. 52–58). Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Duncan, G. J., Dowsett, Ch. J., Claessens, A., Magnuson, K., Huston A. C. et al. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446.

Evans, J. J., Floyd, R. G., McGrew, K. S. and Leforgee, M. H. (2002). The relations between measures of Cattell–Horn–Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review*, *31*(2), 246–262.

Gathercole, S. E., Alloway, T. P., Willis, C. and Adam, A.-M. (2006).Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, *93*(3), 265–281.

Jones, M. and Mandeville, K. (1990). The effect of age at school entry on reading achievement scores among South Carolina students. *Remedial and Special Education*, *11*(2), 56–62.

Kaczan, R. and Rycielski P. (eds.). (2014). B*adanie 6- i 7-latków – rok szkolny 2012/2013*. Warszawa: Instytut Badań Edukacyjnych.

King, G. (1986). How not to lie with statistics: avoiding common mistakes in quantitative political science. *American Journal of Political Science*, *30*(3), 666–687.

Konarzewski, K. (1993). W jakich warunkach inteligencja i stopnie uczniów są współzależne? *Kwartalnik Pedagogiczny*, 1, 111–126.

Konarzewski, K. (2012). *Osiągnięcia szkolne polskich trzecioklasistów w perspektywie międzynarodowej.* Warszawa: Centralna Komisja Egzaminacyjna.

Konarzewski, K. (2013). Wiek startu szkolnego a osiągnięcia w nauce w okresie wczesnoszkolnym. *Edukacja*, *124*(4), 5–19.

Konold, T. R. and Pianta, R. C. (2005). Empirically-derived, person-oriented patterns of school readiness in typically-developing children: description and prediction to first-grade achievement. A*pplied Developmental Science*, *9*(4), 174–187.

Kurdek, L. A. and Sinclair, R. J. (2001). Predicting reading and mathematics achievement in fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology*, *93*(3), 451–55.

La Paro, K. M. and Pianta, R. C. (2000). Predicting children's competence in the early school years: a meta-analytic review. *Review of Educational Research*, *70*(4), 443–484.

Langer, P., Kalk, J. M. and Searls, D. T. (1984). Age of admission and trends in achievement: a comparison of Blacks and Caucasians. *American Educational Research Journal*, *21*(1), 61–78.

Mullis, I. V. S., Martin, M. O., Foy, P. and Arora, A. (2012a). *TIMSS 2011 international results in mathematics.* Chestnut Hill: Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P. and Drucker, K. T. (2012b). *PIRLS 2011 international results in reading.* Chestnut Hill: Boston College.

Mussolin, Ch., Nys, J., Content, A. and Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS ONE*, *9*(3), 1–12.

Oshima, T. C. and Domaleski, Ch. S. (2006). Academic performance gap between summer-birthday and fall-birthday children in grades K-8. *Journal of Educational Research*, *99*(3), 212–217.

Princiotta, D., Flanagan, K. D. and Hausken, E. (2006). *Fifth grade: findings from the fifth-grade follow-up of the early childhood longitudinal study, kindergarten class of 1998–99* (NCES 2006-038). Washington: National Center for Education Statistics.

Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.

Sharp, C. (2002). *School starting age: European policy and recent research*. Paper presented at the LGA Seminar "When should our children start school?", London.

Verachtert, P., De Fraine, B., Onghena, P. and Ghesquière, P. (2010). Season of birth and school success in the early years of primary education. *Oxford Review of Education*, *36*(3), 285–306.

Wilgocka-Okoń, B. (2003). *Gotowość szkolna dzieci sześcioletnich*. Warszawa: Żak.

This article is based on data gathered in research carried out within the systemic project "Quality and effectiveness of education – strengthening of institutional research capabilities" implemented by the Educational Research Institute and co-financed by the European Social Fund (Human Capital Operational Programme 2007–2013, Priority III High quality of the education system). A preliminary version of this article was published in Polish in *Edukacja*, *133*(2), 2015.