Edukacja 2014, 5(130), 5–18
An interdisciplinary approach
ISSN 0239-6858

# The effect of school entrance age on academic performance

Krzysztof Konarzewski

Educational Research Institute*

The study investigated the relationships between school achievement and pupil age at entry into first grade and age relative to the class. Data from 101 519 grade 4 pupils were selected from TIMSS 2011 achievement data in mathematics and science, collected from national samples of 25 European countries. Hierarchical regression analysis showed that the effect of class relative age was greater than the effect of grade relative age and it was significantly higher in younger than older classes. Average achievement (especially in science) was better in older classes.

Keywords: education, academic achievement, birth date effect, relative age effect, TIMSS.

The age children should start school is one of the most common questions posed to education researchers. It is centred on the belief that there is an optimal school start age that maximises the probability of educational success by all or the majority of children and that the answer can be arrived at scientifically. Unfortunately, this is impossible.

Firstly, school starting age has an effect on education in its interaction with the education system: curriculum, teaching organisation, qualifications and professional attitudes of teachers. It is quite possible that an optimal age in one system would prove to be non-optimal in another. Secondly, even with a defined education system, it would not be possible to determine the optimal age, as it would require long-term experiment in which samples of five-, six-, seven- and eight-year-old children chosen at random would start school. Such an experiment is impossible because it is difficult to imagine parental consent to accelerate or delay the school enrolment of their children in relation to the tradition of the country. The results would be adversely affected by awareness of children, parents and teachers being involved in an experiment and interactions of the independent variable with external events, e.g. discrimination of "delayed" pupils.

Researchers can only compare academic achievement[1] in countries differing in age of compulsory school enrolment. Unfortunately, such studies do not generate binding

This article was published primarily in Polish in *Edukacja, 124*(4) 2013.

* Address: Instytut Badań Edukacyjnych, ul. Górczewska 8, 01-180 Warszawa, Poland. Email: k.konarzewski@ibe.edu.pl

[1] Educational achievement is the entirety of cognitive, affective and psychomotor changes in the mind of a student, compatible with the implicit or explicit goals of education and occurring as a result of school attendance. In line with this broad definition, mastering multiplication tables, learning to swim or improving cooperation with others may be regarded as an educational achievement.

conclusions because countries compared, differ not only in the age of compulsory school enrolment, but also in culture, wealth and organisation of the education system, i.e. factors that may influence achievement irrespective of age. A study by Elley (1992, quoted after Sharp, 2002), is a striking example of inconclusiveness of the comparison. Reading comprehension measurements were juxtaposed with school starting age in 32 countries and it was concluded that in the 10 countries with the best results, the age of compulsory school enrolment was slightly higher (6.3 years) than in the 10 countries with the lowest results (5.9 years). However, the difference has reversed when the researcher statistically controlled for the national economic development variable. Probably, this result could be further changed by statistical control of other variables.

More clarification may be provided by comparison of educational achievement of age groups within a grade or class[2], as these groups are educated in similar conditions. In the majority of contemporary education systems, children who meet the age criterion specified in education law start school together on the same day, for example in Poland on the first (working) day of September. This criterion defines the age of compulsory school enrolment[3]. School education is compulsory for every child whose age on the official school enrolment day is not lower than the criterion. For example, in line with the Act on the Education System amended in 2008, a child starts school on the 1st of September in the year when they turn six years old, not later than on the 31st of December, in the Czech Republic and Slovakia, not later than on the 31st of August, and in Hungary, not later than on the 31st of May. In each case, age distribution within first grade is at least 1 year. When differences between younger and older pupils in a class in terms of educational achievement or emotional health measurements are revealed, it is tempting to attribute such differences to age. In the existing literature on the subject these differences are referred to as birth date effects.

### Studies on birth date effects

The first reports on the subject were made in the 1930s (Bigelow, 1934). They were confirmed with subsequent studies. King (1955) measured the achievement of 104 pupils at the end of sixth grade, some of whom started school before and others after reaching the age of six. Educational achievement of older pupils proved to be significantly higher. According to the author, this is explained by greater ability of older children to cope with the stress that school inevitably causes. Similar results were obtained by other researchers (Allen and Barnsley, 1993; Bedard and Dhuey, 2006; Davis, Trimble and Vincent, 1980; Martin, Foels, Clanton and Moon, 2004; Thompson, 1971). It was also found that older children were more often directed to educational programs for gifted pupils (Maddux, Stacy and Scott, 1981) and placed in more advanced streams (Freyman, 1965; Sutton, 1967).

Many studies show that relatively younger children encounter more obstacles to learning and adapting to school life, e.g. they are more likely to need to repeat a year (Dobkin and Ferreira, 2009; Langer, Kalk and Searls, 1984; Verachtert, De Fraine, Onghena and Ghesquière, 2010). DiPasquale and colleagues (1980) established the birth dates of

[2] Grade is the period of education that often, but not always, covers one school year. Class, in Poland erroneously referred to as grade, is a group of students who often, but not always, belong to one grade and maintain their composition in most school subject classes.

[3] In the United States, the criterion of obligation has been replaced with the criterion of eligibility. A child that meets the criterion is entitled to go to school, but the decision on sending a child to school remains with parents. Keeping eligible children in a kindergarten for an additional year is termed academic redshirting, alluding to shirts worn by players out of the game in sports teams (Graue and DiPerna, 2000).

552 pupils in all grades in a school district in the United States who were referred by teachers to psychological counselling as a result of learning difficulties or behavioural problems (71% were boys). The researchers found that the percentage of children referred was linearly dependent on the age when they were enrolled in kindergarten. Thirteen percent of children born in January and February were referred to counselling, compared with 26% of children born in November and December. The effect was significant only for boys in the three initial grades and only those experiencing learning difficulties. In another study however (Weinstein, 1969), the effect was also revealed with behavioural problems. Diamond (1983) and Maddux (1980) found that younger children are more often placed in therapeutic or remedial classes and are less popular with others in the class (Miller and Norris, 1967).

This research trend was soon subject to criticism, especially because some psychologists and educationalists used the concept of school maturity to justify keeping younger children in kindergarten for an extra year (Uphoff and Gilmore, 1986). Gredler (1980), a school psychologist, found that some teachers stereotypically believed that the youngest pupils in a class, especially boys, would cause problems and therefore referred them to psychologists for reasons that they would have ignored in older children, especially girls. It seems that his claim was right, as subsequent studies revealed that the youngest children are more often erroneously diagnosed with learning difficulties (Gledhill, Ford and Goodman, 2002). Gredler was also right, in that children who lacked school maturity needed not so much time, but active help from the education system.

Recent studies, also in the UK, Norway and Canada (Reijneveld et al., 2006), extended the range of dependent variables by including emotional health measurements. In a study in Alberta (Thompson, Barnsley and Dyck, 1999), a higher rate of suicide was found among younger pupils. This effect could not be explained by lack of school maturity alone. A more probable cause seemed to be a combination of lack of school maturity and the social mechanism for comparison of oneself with others. This astonishing fact that in league hockey teams there are more players born in the first than the last months of a given year lead Thompson's (2004) team to the assumption that players are recruited to teams in a similar way to school enrolment, i.e. on the age criterion. As a result, the youngest team players, with less developed psycho-motor abilities are less effective in their roles. Since competition forces them to compare themselves with others, they more often suffer from a sense of inferiority. Justifying this with lack of ability they risk permanent decrease in their self-esteem and reduced self-esteem forces them to leave the team.

The above reasoning includes many assumptions, each of which would require verification. Nevertheless, Thompson and colleagues (2004) apply this reasoning to school, hypothesising that reduced self-esteem in early school years may link the age of compulsory school enrolment with later suicidal tendencies. This is supported by the results of a self-esteem study conducted on 1129 Canadian pupils in grades 1 to 9. Representativeness of the sample is unknown, but the fact that it was composed of more pupils in grades 1–6 (90%) than 7–9 (10%) gives rise to doubt. Self-esteem was measured with two versions (for younger and older children) of the self-esteem inventory by Battle (Culture Free Self-Esteem Inventory, c.f. Brooke, 1995), but they were considered equivalent in the analysis. The results for pupils of different age groups were analysed together, ignoring significant time differences between the experience presumed as the source for self-esteem and its measurement. A weak, but significant linear birth date effect was reported, along with a much stronger (judging from the $F$ value[4]) effect of family structure (pupils from broken homes had lower self-esteem than those from intact families). There was no interaction between the two variables.

Some results of this study are inconsistent with the theory of Thompson and colleagues (2004). They incorporated into the research design two age extreme groups: early starters (enrolled at school 3 months younger than required) and late starters (entering school three-months older than the normal intake). These groups matched the general trend perfectly, contrary to the theory of initial school maturity:

- The children in the first group satisfied the school maturity requirement, otherwise they would not have been accepted. They might be expected to successfully compete with the older children, but their self-esteem was lower. Why?
- Children in the second group benefitted from postponement of school, the practice for those who are not deemed ready for school or who repeated a grade and therefore were delayed. Yet their self-esteem was the highest. Why?

These observations demand some explanation.

In light of these paradoxes and the empirical evidence (Konarzewski, 2013), this attempt to establish a relationship between the school starting age and suicidal tendencies seems flawed.

Many studies have shown that the birth date effect on achievement decreases with age (Dolata and Pokropek, 2012; DiPasquale et al., 1980; Jones and Mandeville, 1990). Langer and colleagues (Langer, Kalk and Searls, 1984) analysed data from a sample from the National Assessment of Educational Progress, including achievement in grades 4 (nine-year-olds), 8 (thirteen-year--olds) and 11 (seventeen-year-olds). Relative and mean ages of children in a grade were considered, separating older classes in states with a September cut-off and younger classes in states with a December cut-off. Stepwise regression analysis with control of sampling design, gender, family socio-economic status and the school environment showed that:

- in the nine-year-old cohort significantly higher achievement was reported for relatively older pupils ($R^2 = 0.2$)[5] and older classes (0.3), irrespective of skin colour,
- in the thirteen-year-old cohort the advantage of relatively older pupils among white pupils decreased (0.1), but it was unchanged among black pupils (0.3); advantage of older classes disappeared for white pupils, but it was unchanged for black pupils,
- in the seventeen-year-old cohort both effects disappeared, both for white and black pupils.

It was also found that younger pupils in the nine- and thirteen-year-old cohort repeated a grade more often than older ones and specifically the white nine-year-old cohort. Younger boys were more likely to repeat a grade than younger girls. The authors conclude that "Successful student adaptation to the school environment and student retention are two of the possible reasons for the decreasing importance of relative age" (Langer et al., 1984, p. 73). The distinction between pupil and pupil class age has not proved a fruitful concept: there was no evidence that the interaction of these two variables controlled any dependent variable. It should be however noted that the interaction was studied simply using the product of both dependent variables. This method

[4] The text (p. 316) mentions only the $F$ values (2.49 and 20.36 respectively). Incidentally, the number of degrees of freedom in the numerator is 1, which is incomprehensible in light of the fact that the analysis covered 6 age groups.

[5] All $R^2$ values reported in the text are multiplied by 100, so that they mean a percent of the explained variance of dependent variable.

ignores differences in local (class) achievement distributions which may mask the interaction.

The majority of birth date researchers used a cross-sectional study scheme, but longitudinal studies were also used. In Poland such a study was conducted by Dolata and Pokropek (2012) on data from the Central Examination Board. The data was from three age cohorts born in the years 1994–1996 and within the age range appropriate for their grade. Regression analysis of achievement twice showed weak birth date effects. Effects on achievement at the end of primary and lower secondary schools were significant but weak in primary pupils (with $R^2$ ranging from 0.29 to 0.38 in three cohorts) and weaker in these same pupils at secondary school, in particular on achievement in mathematics and science (0.03–0.07). This reasonably confirmed Langer's results (1984). A British study (Hutchison and Sharp, 1999, quoted in Sharp, 2002), using the same scheme as Dolata and Pokropek (2012), measured reading comprehension. Older pupils scored higher but the difference decreased with time in education from 0.47 to 0.25 (Cohen's *d* units) in twelve-year-olds.

Work by Verachtert et al. (2010) was based on data from a mass longitudinal study of mathematical achievement by 3990 pupils from 122 schools in Belgium (3156 children from 120 schools participated in the last measurement). Achievement was measured at three times: the beginning of grade 1 and the ends of grade 1 and grade 2, using vertically aligned tests. These tests allowed quantification of progress in mathematics using multilevel modelling of growth curves. The level of individualised instruction was taken into account, using two indicators: the relative amount of "frontal" teaching time, i.e. addressed to the entire class and frequency of teaching addressed to children with the lowest and the highest achievement. Both indicators were based on a teacher survey, which can be a rather unreliable source. The study results may be summarised in a few points.

- At the beginning of grade 1, the children from the relevant age group born in the first quarter (the oldest ones) achieved significantly higher mathematics test scores than children born in the fourth quarter (the youngest ones) ($d = 0.43$), but for over-age children the results were opposite: the younger children showed an advantage over the older ones. The authors explain the opposite effect using different reasons for postponement of compulsory school enrolment: in younger children it would probably be transient immaturity, in older children – permanent lesser ability.
- The percentage of children within the appropriate age range, who did not advance to grade 2, increased from 6.4 for the children born in the first quarter to 20.0 for children born in the last quarter.
- When compared with children born in the first quarter, the children born in the second half of that year were developing more rapidly, so that the initial gap in achievement was reduced by half[6]. This effect also applies to delayed pupils (who repeated kindergarten in the past).

The "frontal" teaching indicator significantly differentiated the pace of children's progress, but to such a small extent ($d = 0.02$) that the authors rightfully wrote: "Maybe, providing autumn-born children with appropriate levels of instruction is not a good way to tackle the season of birth effect in education" (Verachtert et al., 2010, p. 303).

An interesting complication of the birth date effect was discovered in economics. Angrist and Krueger (1991) showed that a smaller percentage of older pupils (born in

[6] However, in a similar, though quantitatively modest study, Morrison et al. (1997) found that during grade 1, the increase in achievements of younger and older pupils in reading and mathematics was equal. Apparently, learning pace depended on teaching strategies rather than student age.

the first quarter) than younger ones (born in the fourth quarter of the previous year) complete secondary school. Dobkin and Ferreira (2009), having analysed data on people over thirty-years-old from a census in California and Texas, concluded that those born just before the cut-off date (i.e. the youngest) more often repeated a grade than those born at least 180 days prior to that date (the difference was 20 percentage points). They also confirmed the result by Angrist and Krueger (1991): slightly more younger people graduated from high school (difference of 1 percentage point). Therefore the youngest ones perform somewhat worse at school, but they are slightly more likely to acquire secondary education. However, the most important, at least from the point of view of economics, was the finding that the relative school enrolment age had no influence on employment, remuneration, home ownership, marital status, family income, etc. in adulthood. This result applies to all categories of gender, age and ethnicity.

Finally, it should be mentioned that in some studies the birth date effect was not found (Black, Devereux and Salvanés, 2008, Dietz and Wilson, 1985; May and Welch, 1986), or was considered practically negligible (Shepard and Smith, 1985, quoted in Morrison et al., 1997), especially when compared to socio-economic factors (Bickel, 1991, Jones and Mandeville, 1990), or school maturity measurements (Wood, Powell and Knight, 1984). Some researchers also reported the opposite effect: relatively younger children performed better in a test of school abilities than older children (McDonald, 2001), and due to postponement of school enrolment, the oldest children in a class exhibited more behavioural problems in adolescence and were more likely to use special education services (Byrd, Weitzman and Auinger, 1997). May and colleagues (1995) believe that postponement of school often conceals permanent learning difficulties labelled as immaturity, so that help for a child is delayed.

## Birth date effect theory

A satisfactory explanation of the birth date effect must refer to the interaction between pupil competence and challenges embedded in the school environment. The current competence level of a pupil is best understood as the ultimate on the scale of educational challenges that the pupil is able to meet à vista, without the need to learn. The main assertion of the theory is: development, i.e. the long-term growth of subject competence, is a curved function of environmental challenge (Figure 1). This means that the scale of challenge includes the optimum $w$ value representing the level that efficiently promotes the development of a given person. Points to the left of $w$ represent insuficient challenge and points to the right are excessive. Nevertheless, even sub-optimal challenge has some developmental value which decreases as the distance from of the optimum increases. It is also assumed that values for challenge which are too large and too small do not differ from each other.

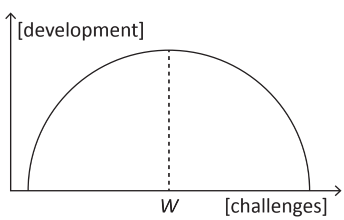


Figure 1. Development as a function of challenges.

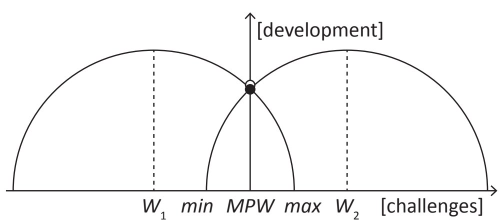


Figure 2. A two-child class.

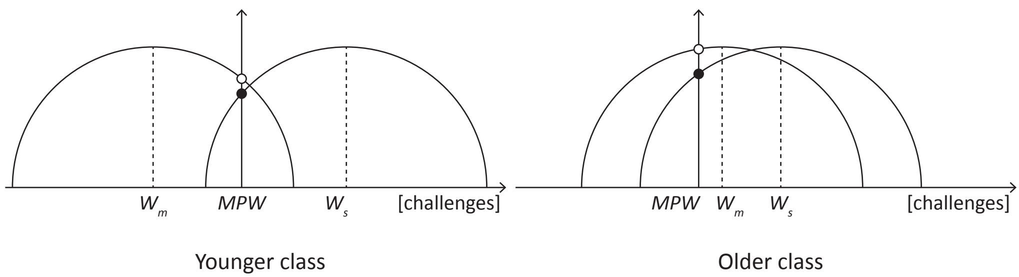


Figure 3. Two two-child classes.

A class may be represented as a nest of individual development curves located on a common scale of challenges with an established modal level (MPW). MPW is determined by both teacher and teaching culture as an aggregate of curriculum requirements, methods of transferring educational content, teaching pace and evaluation style. In countries where the cult of childhood dominates (recognition of childhood as a distinct form of humanity, requiring special protection from the adult world), MPW is shifted to the left, towards lower challenge. It is likely that in these countries, there is a tendency to raise school enrolment age.

Figure 2 shows a class consisting of two pupils with different levels of competence. The difference between the individual *w* points expresses the intra-class competence diversity. In the figure, MPW is optimally located: at the intersection of the two curves, which means that the class environment stimulates the development of both pupils equally and at the same time, maximises average development for the class. MPW shift to the left would favour pupils with lower competences and a shift to the right, pupils with higher competences. The permissible range of MPW variation defines the minimum and maximum values. Exceeding the range implies that one pupil ceases to make progress.

Competence essential for school life of children entering school is commonly known as school maturity or school readiness (Ilg and Ames, 1965). The level of this competence depends on many factors. One such factor is pupil age. May and Welch (1986) demonstrated that the results of the popular maturity test (Gesell School Readiness Screening) are associated with children's date of birth and that this relationship fades with age[7]. The reasons for these effects are subject to strong debate (c.f. Lawlor, Clark, Ronalds and Leon, 2006; Martin et al., 2004) not referred to in this article. What is most important is the difference in initial competence that leads to lower expected achievement from younger children, which is higher in younger classes (i.e. with pupils of a lower average age) than in older classes.

Figure 3 shows two two-child classes: a younger class with a lower mean age and an older class with a higher mean age. In the younger class, the initial difference between the curves is greater, the permissible MPW range is narrower and the MPW is located near the intersection of the curves, which means that both the younger and older pupils develop at a similar rate. In the older class, the differences between pupils are smaller, the permissible MPW range is wider and the MPW may favour the younger pupil, so that they develop at a higher rate than the older, but both achieve more than the younger class.

What can we expect in the younger and older classes after a few years of study, for

[7] However, the relationship between school maturity at the start of schooling and subsequent achievement and school adaptation sometimes would be found in the data (Banerji, 1992; Graue and Shepard, 1989), and sometimes would not (de Lemos and Mellor, 1994).

example in grade 4? The theory leads to three hypotheses:

- $H_1$: achievement of younger pupils in a class will be lower than those of older pupils;
- $H_2$: in younger classes achievement differences between younger and older pupils will be greater than in older classes;
- $H_3$: in younger classes the achievement average is lower than in older classes.

These three hypotheses were tested.

## Methodology

The analysis was conducted on the data from the international IEA TIMSS 2011 study (Konarzewski, 2012, Martin, Mullis, Foy and Stanco, 2012, Mullis, Martin, Foy and Arora, 2012). The aim of the study was to determine educational achievement in mathematics and science of ten-year-olds in grade 4 from 50 countries. It was not possible to apply both conditions for the international definition of population in all countries – in some, the fourth-graders were younger (e.g. Italy) or older (e.g. Denmark) than 10 years of age, in other countries ten-year-olds studied in a higher grade (e.g. England). The measurement was carried out in May 2011. In Poland the study covered third-grade pupils with a mean age of 9.9 years.

The database containing the data of over 600 000 pupils was reduced in three steps. Firstly, in order to reduce the number of organisational variants, all countries outside of Europe were eliminated, leaving data from 25 European countries. Secondly, "delayed" pupils, i.e. older than appropriate for their grade, were excluded. Some of these pupils, as shown in the above studies, come from a different population than pupils of an age appropriate for their grade. Many "delayed" pupils have experienced and still experience learning difficulties, which influences their achievement more than age, therefore including them in the analysis could distort the relationship between age and achievement. For this reason, from the initial number of 108 392 pupils, 5043 (4.7 %) "delayed" pupils were excluded – most from Germany (12.7%) and Austria (10.6 %), – the fewest from Norway (0.5%), Croatia (0,9%) and Poland (1.9%).There is no need, however, to exclude "advanced" pupils from the analysis. They went to school before the age of compulsory school enrolment, but they were mature enough to learn, otherwise they would not have been accepted early. Thirdly, atypical classes with fewer than 5 pupils or more than 32 pupils were excluded. These accounted for 3.9% of all classes. Finally, the analysis covered 101 519 pupils from 5585 classes with a mean class size of 18.2 with standard deviation of 5.6. The age of pupils ranged from 6.3 to 11.6 years around the mean 10.3 with standard deviation of 0.51. The sample was 50.3 % boys.

Analysis of school achievement requires control of confounding variables – at least gender and socioeconomic status (SES) of pupils' families. In the IEA studies, the SES indicator is not created, but data from which it may be generated are collected. The indicator was generated separately for each country as a factor score (c.f. Konarzewski, 2012, p. 64), hence national distributions do not differ[8].

The analysis was performed using the two-level hierarchical linear model (Raudenbush and Bryk, 2002). The great advantage of this is that it allows estimation of the interesting dependencies on the site where they arise. In this case, birth date effect arises in a class, therefore it is there that the relationship between pupils' achievement and age should be determined, at its simplest as the slope in the

[8] In five countries (England, Belgium, Denmark, Netherlands, and Serbia) parents did not fill in the questionnaire and in other countries calculating SES was difficult due to missing data. Overall, SES was unknown for 19% of the pupils. In these cases an evaluation of SES has been used, based on the included in the SES definition indicators of wealth from the pupil's questionnaire.

class regression equation. In a strict sense, this should be used to examine relative age effect rather than birth date effect. A pupil in the middle of the grade age distribution may be classified to the group of youngest or oldest children in the class. If relative age is more important than absolute age, its effect should be stronger than the birth date effect.

Five plausible values for pupil achievement in mathematics and science were the dependent variables. At the first level of regression analysis, an attempt was made to explain them using the family status, gender and age. All independent variables were "centered", i.e. expressed as deviations from the average in the class. This allowed the regression equation constant to be equal to the mean class achievement. At the second level, an attempt was made to explain class achievement averages and slopes, using mean age of pupils in each class.

## Results

Results of the regression analysis are presented in Tables 1 and 2. It may be observed that SES and gender are significantly correlated with achievement, if they were not included in the equation, the dependence would be distorted as a result of random fluctuations in the classs composition.

The intra-class mathematical achievement regression coefficients on relative age are highly varied ($u_3$), but their average is positive and significantly greater than zero ($\gamma_{30}$). In an average class, a one year difference translates into a difference of 5.93 points. By dividing it by the standard deviation of the dependent variable (61.12), we obtain $d = 0.10$. The relationship between the relative age and the achievement in science is similar: $d = 0.14$. The percentage strength of the effect calculated by comparing the effect of variance explained by the model including the relative age variable and not including it, is 1.27 for mathematics and 1.26 for natural sciences, i.e. over six times the strength of the birth date effect detected in the nine-year-olds group by Langer's (1984) team and almost four times higher than the effect in the twelve-year-olds group (Dolata and Pokropek, 2012). Hypothesis $H_1$ was confirmed.

Table 1
*Fixed effect estimates with robust standard errors*

| | Mathematics | | | | Science | | | |
|---|---|---|---|---|---|---|---|---|
| Effects | Coefficient | Standard error | *df* | *p* | Coefficient | Standard error | *df* | *p* |
| Achievement | | | | | | | | |
| Average [$\gamma_{00}$] | 521.82 | 1.30 | 352 | <0.001 | 530.00 | 1.43 | 35 | <0.001 |
| Average age in a class [$\gamma_{01}$] | 25.68 | 3.01 | 3 085 | <0.001 | 23.93 | 2.89 | 2 954 | <0.001 |
| Age | | | | | | | | |
| Slope [$\gamma_{30}$] | 5.93 | 1.22 | 3 281 | <0.001 | 8.16 | 1.37 | 99 | <0.001 |
| Average age in a class [$\gamma_{31}$] | -10.58 | 2.97 | 402 | <0.001 | -10.24 | 3.79 | 26 | 0.012 |
| SES slope [$\gamma_{12}$] | 15.19 | 0.50 | 304 | <0.001 | 16.14 | 0.55 | 49 | <0.001 |
| Gender slope [$\gamma_{23}$] | 6.49 | 0.80 | 274 | <0.001 | 5.84 | 0.90 | 53 | <0.001 |

The average age of pupils in a class ($\gamma_{31}$) negatively differentiates the intra-class achievement regression coefficients on relative age. This means that the older the children in a class, the lower the advantage of relatively older pupils over the younger ones. This confirms the $H_2$ hypothesis.

The average age of pupils in a class ($\gamma_{01}$) is positively associated with the achievement average in the class. Achievement of pupils in the older classes was higher in the younger children classs, which is in line with hypothesis $H_3$.

Table 3 shows the same results in a more intuitive way. Classes were divided into three equal subsets based on the mean absolute age, in six-month intervals. In each subset, intra-class achievement regression coefficients on relative age were determined. It may be observed that the differences related to relative age (expressed as regression coefficients) are the highest in the youngest classes, smaller in medium-aged classes and indistinguishable from zero in the oldest classes. Secondly, contrary to equality of the achievement regression coefficients on average age ($\gamma_{01}$ in Table 1), this variable differentiates the average mathematical achievement other than science achievement. In science, achievement was higher, the older the children in a class. Pupils the in medium-aged classes (i.e. those who started school at age 6.6) were more able at mathematics than pupils in the younger classes (who went to school six months earlier), but further postponement of school enrolment was without benefit: pupils in the older classes (who started school at age 7) did not achieve more than those in classes of children six months younger. Thirdly, the older the children in a class, the less the

Table 2
*Variance components estimates*

| Effects | Mathematics | | | Science | | |
|---|---|---|---|---|---|---|
| | Variance | $\chi^2$ *df* ≈ 5583 | *p* | Variance | $\chi^2$ *df* ≈ 5583 | *p* |
| Average [$u_0$] | 1 911.10 | 54 157 | <0.001 | 1 650.83 | 47 006 | <0.001 |
| Age slope [$u_3$] | 341.79 | 6 613 | <0.001 | 315.14 | 6 590 | <0.001 |
| Level 1 [*r*] | 3 735.72 | — | — | 3 789.03 | — | — |

Estimated model: $Y_{ij} = \gamma_{00} + \gamma_{01}\, Average\_age_j + \gamma_{10}\, SES_{ij} + \gamma_{20}\, Gender_{ij} + \gamma_{30}\, Age_{ij} + \gamma_{31}\, Average\_age_j\, Age_{ij} + u_{0j} + u_{3j}\, Age_{ij} + r_{ij}$. The data at level 1 were weighted with the population weight.

Table 3
*Educational achievement (controlling for gender and SES)*

| Class | Number of classes | Average age distribution | Average age | Mathematics: Grand mean achievement | Mathematics: Slope of achievement regression on age – Value | Mathematics: Slope of achievement regression on age – *p* | Science: Grand mean achievement | Science: Slope of achievement regression on age – Value | Science: Slope of achievement regression on age – *p* |
|---|---|---|---|---|---|---|---|---|---|
| Younger | 1 857 | 9.36–10.06 | 9.83 | 497.9 | 9.99 | <0.001 | 511.8 | 12.99 | <0.001 |
| Medium | 1 868 | 10.07–10.58 | 10.31 | 528.7 | 6.48 | 0.003 | 528.4 | 7.44 | <0.001 |
| Older | 1 860 | 10.59–11.22 | 10.78 | 526.9 | 2.12 | *ns* | 540.4 | 4.74 | *ns* |

difference in achievement was between boys and girls. In the youngest classes, the difference in achievement in mathematics and science was 9.6 and 7.9 points in favour of boys and in the oldest classes, significantly lower: 3.1 and 2.1.

## Discussion

Confirmation of $H_1$ may seem trivial in the light of the many studies in which this effect has been shown. A particular feature of this study however should be noted. The relative age in the sample of classes from different education systems depends on the season of birth to a minor extent. In fact, the youngest and oldest pupils in their classes were born in all months of the year: the oldest most often in January (in 31% of classes), and least often in April (4%), the youngest most often in December (30%) and least often in April (5%). The relative age effect is therefore of a particularly school-related nature and cannot be reduced to biological and weather conditions during fetal development and early infancy.

Confirmation of $H_2$ means that there is an interaction, which Langer and colleagues (1984) failed to detect, between the relative and absolute school enrolment age. Postponement of school enrolment by one year cancels the relative age effect probably because it allows a teacher to better adapt to the challenging needs of younger pupils.

$H_3$ hypothesis has only been fully confirmed with respect to science. Why does not mathematical achievement in the older classes differ from classes which are on average 6 months younger? If in systems of late school enrolment the cult of childhood is stronger, MPW shift to the left, towards less demanding challenges might be expected. Relatively younger children catch up more quickly with the older ones, but meet them at a lower level than would be expected if requirements were higher for everyone. This only observed in mathematics classes since mathematics is considered to be particularly difficult. The discrepancy of the results from those that would be expected according to the $H_3$ hypothesis, suggests that MPW for different subjects needs to be analysed separately.

The study also produced an unexpected effect: in the older classes the advantage of boys over girls is lower than in the younger classes. The theory provides a simple explanation. At the start of school education boys have an advantage over girls in terms of numeracy skills[9]. In order to account for this difference, it is necessary to divide every function in Figure 3 according to gender; girls (shifted slightly to the left) and boys (shifted slightly to the right). It seems obvious that older classes provide more stimulation for girls than boys, while younger classes provide roughly equal amounts.

***

How relevant are these results to Polish debate about school enrolment age? The findings are not directly conclusive, but legitimise the claim that age relates to school achievement through initial aptitude differences and teaching strategies. The earlier children start school, the more diversity a teacher is facing and the harder it is for the teacher to pitch the level of challenge that would maximise pupil achievement. How to reduce those differences is a more important question than defining the age at which children start school.

For experts on the process of maturation who believe that a child, like wine, matures exclusively due to the passage of time, the answer is simple: it is necessary to postpone enrolment. However, it is clear that what matters is not simply time, but how it

[9] No research was found to confirm this directly, but sound British research (Calvin et al., 2010) shows the advantage in the cohort of eleven-year-olds. It is likely that the advantage is no less in the cohort of six-year-olds.

is spent. If children were to spend an extra year at home, differences in grade 1 would further increase, since domestic developmental challenges are closely related to family socioeconomic status. If children attend "protoschool" (in Poland, commonly known as reception grade), i.e. school without explicit demands but with continuous supervision over their activity and consequent restriction and redirection, their "proto--start" takes place in an even more diverse group than the actual start, which should reinforce the relative age effect. The most effective solution seems to be one year spent in kindergarten that does not teach specifically school-related skills, but provides appropriate development stimulation. Such a year could be offered to both five- and six-year-olds.

Apart from improvement of kindergarten education, differences at the beginning of grade 1 may be reduced further by adjusting the cut-off date. Estimated school enrolment age for pupils in the youngest classes (Table 3) is 6.2 years, in the medium-aged classes – 6.6 years and for those in the oldest classes – 7.1 years. These values correspond to the following cut-offs: age 6 and 31st December, age 6 and 30th June and age 7 and 31st December. The latter situation existed in Poland before the legal amendment; the former will apply when the amendment comes into force. The compromise: enrolling those children who turn six before the 30th June would reduce the relative age effect without subverting the intended effect of the change.

Two other possible solutions might assign pupils to classes and to the reorganisation of education in classes. The first assumes creation of separate classs for younger and older children. The effectiveness of this solution would need to be tested by experiment. The second solution assumes a more profound individualisation of early school instruction, especially in line with the Joplin plan (c.f. Konarzewski, 2011).

## Literature


Allen, J. and Barnsley, R. H. (1993). Streams and tiers: the interaction of ability, maturity and training in systems with age-dependent recursive selection. *Journal of Human Resources*, *28*, 649–659.

Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, *106*(4), 979–1014.

Banerji, M. (1992). An integrated study of the predictive properties of the Gesell School Readiness Screening Test. *Journal of Psychoeducational Assessment*, *10*(3), 240–256.

Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: international evidence of long-run age effects. *Quarterly Journal of Economics*, *121*(4), 1437–1472.

Bickel, D. D. (1991). Chronological age and entrance to first grade: effects on elementary school success. *Early Childhood Research Quarterly*, *6*(2), 105–117.

Bigelow, E. B. (1934). School progress of under-age children. *Elementary School Journal*, *25*, 186–192.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2008). Too young to leave the nest: the effects of school starting age. NBER Working Papers No. 13969.

Brooke, S. L. (1995). Critical analysis of the culture-free self-esteem inventories. *Measurement & Evaluation in Counseling & Development*, *27*(4), 248–252.

Byrd, R. S., Weitzman, M. and Auinger, P. (1997). Increased behavior problems associated with delayed school entry and delayed school progress. *Pediatrics*, *100*(4), 654–661.

Davis, B. G., Trimble, C. S. and Vincent, D. R. (1980). Does age of entrance affect school achievement? *The Elementary School Journal*, *80*, 133–143.

Diamond, G. H. (1983). The birth date effect? A maturational effect? *Journal of Learning Disabilities*, *16*, 161–164.

Dietz, C. and Wilson, B. J. (1985). Beginning school age and academic achievement. *Psychology in the Schools 22*(1), 93–94.

DiPasquale, G. W., Moule, A. D. and Flewelling, R. W. (1980). The birth date effect. *Journal of Learning Disabilities*, *13*, 4–8.

Dobkin, C. and Ferreira, F. (2009). Do school entry laws affect educational attainment and labor market outcomes? NBER Working Paper No. 14945.

Dolata, R. and Pokropek, A. (2012). Czy warto urodzić się w styczniu? Wiek biologiczny a wyniki egzaminacyjne [Is it worth it to be born in January? Biological age and examination results]. In B. Niemierko and M. K. Szmigel (eds.), *Regionalne i lokalne diagnozy edukacyjne*. Materiały XVIII Konferencji Diagnostyki Edukacyjnej [R*egional and local educational diagnoses*. Materials of XVIII Conference on Educational Diagnostics] (pp. 52––58). Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Freyman, R. (1965). Further evidence on the effect of date of birth on subsequent school performance. *Educational Research*, *8*, 58–64.

Gledhill, J., Ford, T. and Goodman, R. (2002). Does season of birth matter? The relationship between age within the school year (season of birth) and educational difficulties among a representative general population sample of children and adolescents (aged 5–15) in Great Britain. *Research in Education*, *68*, 41–47.

Graue, M. E. and DiPerna, J. (2000). Redshirting and early retention: who gets the 'gift of time' and what are its outcomes? *American Educational Research Journal*, *37*, 509–534.

Graue, M. E. and Shepard, L. A. (1989). Predictive validity of the Gesell School Readiness Tests. *Early Childhood Research Quarterly*, *4*(3), 303–315.

Gredler G. R. (1980). The birth date effect: fact or artifact? *Journal of Learning Disabilities*, *13*(5), 9–12.

Ilg, F. L. and Ames, L. B. (1965). *School readiness*. New York: Harper & Row.

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

King, I. B. (1955). Effect of age of entrance into grade I upon achievement in elementary school. *The Elementary School Journal*, *55*, 331–336.

Konarzewski, K. (2012a). *Osiągnięcia szkolne polskich trzecioklasistów w perspektywie międzynarodowej* [*School achievement of Polish third-graders in an international perspective*]. Warszawa: Centralna Komisja Egzaminacyjna.

Konarzewski, K. (2012b). *Perspektywy indywidualizacji kształcenia. Raport o stanie badań* [*Perspectives on individualized instruction. Report on the state of research*]. Warszawa: Instytut Badań Edukacyjnych.

Konarzewski, K. (2013). *Względny wiek uczniów w oddziale a ich samoocena*. Referat przedstawiony na XIX Krajowej Konferencji Diagnostyki Edukacyjnej [*The relative age of pupils in the classroom and their self-esteem*. Paper presented at the XIX Conference on Educational Diagnostics], Gniezno.

Lawlor, D. A., Clark, H., Ronalds, G. and Leon, D. A. (2006). Season of birth and childhood intelligence: findings from the Aberdeen Children of the 1950s cohort study. *British Journal of Educational Psychology*, *76*, 481–499.

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.

Lemos, M. M. de and Mellor, E. J. (1994). A longitudinal study of developmental maturity, school entry age, and school progress. *Australian Research in Early Childhood Education*, *1*, 42–50.

Maddux, C. D. (1980). First-Grade entry age in a sample of children labeled Learning Disabled. *Learning Disability Quarterly, 3*(2), 79–83.

Maddux, C. D., Stacy, D. and Scott, M. (1981). School entry age in a group of gifted children. *Gifted Children Quarterly*, *15*, 180–184.

Martin, R. P., Foels, P., Clanton, G. and Moon, K. (2004). Season of birth is related to child retention rates, achievement, and rate of diagnosis of specific LD. *Journal of Learning Disabilities*, *37*, 307–317.

Martin, M. O., Mullis, I. V. S., Foy, P. and Stanco, G. M. (2012). *TIMSS 2011 international results in science*. Chestnut Hill: Boston College.

May, D. C., Kundert, D. K. and Brent, D. (1995). Does delayed school entry reduce later grade retentions and use of special education services? *Remedial and Special Education*, *16*(5), 288–294.

May, D. C. and Welch, E. (1986). Screening for school readiness – the influence of birth date and sex. *Psychology in the Schools*, *23*(1), 100–105.

McDonald, G. (2001). Comparing school systems to explain enduring birth date effects. *Compare*, *31*, 381–391.

Miller, W. and Norris, R. C. (1967). Entrance age and school success. *Journal of School Psychology*, *6*, 47–60.

Morrison, F. J., Griffith, E. M. and Alberts, D. M. (1997). Nature-nurture in the classroom: entrance age, school readiness, and learning in children. *Developmental Psychology*, *33*(2), 254–262.

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

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

Reijneveld, S. A., Wiefferink, C. H., Brugman, E., Verhulst, F. C., Verloove-Vanhorick, S. P. and Paulussen, T. G. W. (2006). Continuous admission to primary school and mental health problems. *BMC Public Health*, *6*, 145. doi:10.1186/1471-2458-6-145.

Sharp, C. (2002). School starting age: European policy and recent research. Paper presented at the LGA Seminar "When Should Our Children Start School?", London.

Sutton, P. (1967). Correlation between streaming and season of birth in secondary schools. *British Journal of Educational Psychology*, *37*, 300–304.

Thompson, D. (1971). Season of birth and success in the secondary school. *Educational Research*, *14*, 56–60.

Thompson, A. H., Barnsley, R. H. and Battle J. (2004). The relative age effect and the development of self-esteem. *Educational Research*, *46*(3), 313–320.

Thompson, A. H., Barnsley, R. H. and Dyck, R. J. (1999). A new factor in youth suicide: the relative age effect. *Canadian Journal of Psychiatry*, *44*, 82–85.

Uphoff, J. and Gilmore, J. (1986). The birth date effect and college athletic participation: some comparisons. *Journal of Sport Behaviour*, *15*, 227–238.

Verachtert, P., Fraine De, 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.

Weinstein, L. (1969). School entrance age and adjustment. *Journal of School Psychology*, *7*, 20–28.

Wood, Ch., Powell, S. and Knight, R. (1984). Predicting school readiness: the validity of developmental age. *Journal of Learning Disabilities*, *17*(1), 8–11.