The Presence of Matthew Effects in Dutch Primary Education, Development of Language Skills Over a Six-Year Period

Hans Luyten and Gerdy ten Bruggencate

J Learn Disabil 2011 44: 444 originally published online 19 July 2011
DOI: 10.1177/0022219411410289

Hammill Institute on Disabilities

Journal of Learning Disabilities
44(5) 444–458

DOI: 10.1177/0022219411410289
http://journaloflearningdisabilities.sagepub.com
SAGE

The Presence of Matthew Effects in Dutch Primary Education, Development of Language Skills Over a Six-Year Period

Hans Luyten[1] **and Gerdy ten Bruggencate**[1]

Abstract

Using a nationally representative sample of 5,150 Dutch students who have been followed over a 6-year period, the presence of the Matthew effect was investigated for general language skills. The analyses do not reveal unmistakable evidence for the supposition that the rich get richer and the poor poorer. On the contrary, in schools with low starting levels students make more progress than in schools with higher starting levels. On the other hand, the analyses do show a widening gap between students with well educated and poorly educated parents. The gap between delayed and accelerated students was found to increase as well, but the initial disadvantage of boys was found to disappear.

Keywords

Matthew effect; reading skills; longitudinal research

The term “Matthew effect” originates from the sociology of science and was coined by Robert K. Merton (1968) to describe the career development of academic personnel. Eminent scientists and scholars are believed to receive more credit for their work than unknown colleagues, even if the quality of their contributions is similar. In a general sense, the term denotes the phenomenon that “the rich get richer and the poor get poorer,” which may be observed in a wide variety of contexts. Its name refers to a quote from the Bible: “For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath” (Matthew 25:29).

The Matthew effect was introduced into the educational field by Walberg and Tsai (1983), who used the term to describe the relation between socioeconomic status and learning achievement. Stanovich (1986) has used the notion of decisively initial advantages more extensively in his conceptualization of the reciprocal processes through which young children acquire reading skills. He points out that to read children need to recognize some words from a text. The better children are able to read, the easier it becomes for them to derive the meaning of new and unknown words from a text. As a result, word knowledge increases and reading becomes easier, which would set a new cycle of expanding word knowledge in motion. Thus, early success in acquiring reading skills is believed to enhance the further growth in reading abilities. In the higher grades, when reading skills are a precondition for further learning, reading difficulties are likely to produce problems in other subjects as well. In addition, the progress of low-achieving students may be hampered because teachers adapt their goals to the performance level of students. In that case, teachers make fewer demands on poor performers. Furthermore, the average achievement level of a student’s classmates may have a stimulating effect on his or her cognitive development. Unfortunately, students with a relatively low achievement level often attend schools with a modest average achievement level. The socioeconomic composition of schools tends to reflect that of the communities in which they are located. Although the relation between socioeconomic background and academic achievement is far from deterministic and varies substantially across societies (e.g., Organisation for Economic Co-operation and Development [OECD], 2004), the overall picture definitely points to lower average achievement levels of students from disadvantaged backgrounds.

The general message of Stanovich’s argument is fairly depressing for students who start their school career with some disadvantage. It seems that they are bound to lose ever more ground to their more fortunate peers. In a prior study

[1]University of Twente, Enschede, Netherlands

Corresponding Author:
Hans Luyten, University of Twente, Faculty of Behavioral Sciences, Department of Educational Organisation and Management, P.O. Box 217, 7500 AE Enschede, Netherlands
Email: j.w.luyten@utwente.nl

on the presence of Matthew effects in Dutch primary education (Luyten, Creemers-van Wees, & Bosker, 2003), the authors pointed to the fact that a poor starting position offers—in principle—much more opportunity for improvement than an initial advantage. In contrast to the Matthew effect, this may be termed the "bigger fish" effect. This label is based on a quotation from the Dutch comic *De Bovenbazen* (The Bigger Fish), by Maarten Toonder (1983):

> The world is unevenly divided: some have nothing and others have everything. When you don't have anything it is possible to gain more—for anyone who can do that, then life is actually fun. But anyone who has everything is no longer happy when he receives something. Instead, he must always be afraid that he loses something for that is the only possibility that remains for him. (p. 157)

The analyses by Luyten et al. (2003), which cover a 4-year period (ages 8 to 12) and relate to language, mathematics, and nonverbal intelligence, reveal examples of both Matthew effects and bigger fish effects. The present article relates to an analysis of a similar but more recent cohort of Dutch students. The study covers a longer period (6 years) and focuses exclusively on language skills. The next section discusses the findings from prior research on the presence of Matthew effects.

The Matthew Effect in Prior Research

Even though the Matthew effect is repeatedly referred to by various authors in the field of education, studies that actually investigated its presence appear to be remarkably rare. Empirical evidence would require data on individual growth trajectories that cover a substantial period in time. Shaywitz et al. (1995) argue that the educational research conducted before 1995 does not provide clear empirical evidence for the existence of a Matthew effect. From the longitudinal studies conducted before 1995 (Clay, 1979; Juel, 1988; Lundberg, 1984), it can be concluded only that students tend to remain at their original, relative performance levels. Since 1995 a number of longitudinal studies have been published that focus specifically on the presence of Matthew effects (Aarnoutse & Van Leeuwe, 2000; Bast, 1995; Luyten et al., 2003; Phillips, Norris, Osmond, & Maynard, 2002; Scarborough & Parker, 2003; Shaywitz et al., 1995).

The study by Shaywitz et al. (1995) describes the progress made by 396 students in Connecticut who were followed for 7 years from Grade 1 (age 6) through Grade 6 (age 12). Their report focuses on cognitive ability as assessed by the *Wechsler Intelligence Scales for Children–Revised* (1974) and reading decoding skills as measured by the reading cluster of the *Woodcock–Johnson Psycho-Educational Battery* (1977). Their analyses reveal a modest Matthew effect for IQ but not for reading decoding skills.

The study by Bast (1995) focuses not only on the presence of the Matthew effect but also on the mechanisms that may create differences in reading progress. This study relates to 235 Dutch students whose reading skills were measured at six occasions, starting at the end of Grade 2 (age 6) until the end of Grade 5 (age 9). Several aspects of reading ability were assessed (vocabulary, deciphering skills, and reading comprehension). A Matthew effect was found only for deciphering skills, but several reciprocal relationships that are in line with Stanovich's framework could be observed. Reading skills and reading attitudes were found to be mutually reinforcing. The same was the case for vocabulary and reading fluency and for phonological and deciphering skills (Bast & Reitsma, 1997, 1998).

The work by Aarnoutse and Van Leeuwe (2000) also relates to Dutch primary education. The development in reading skills of 556 students was tracked from Grade 3 (age 7) to Grade 8 (age 12). This study produced findings only contradictory to the Matthew effect. For three out of four reading skills (reading comprehension, vocabulary, and spelling) it was found that the initially poor readers actually made more progress. Only for word recognition did the authors fail to find such a "bigger fish" effect.

The study by Scarborough and Parker (2003) relates to a sample of 57 New Jersey students. Their scores on reading achievement, IQ, and behavior problems in second (age 8) and eighth grade (age 14) were assessed. Changes over time were compared for students with no learning disabilities versus students who had been diagnosed as having math or reading disabilities. The findings showed a widening of differences in IQ between students with and without learning problems in math. Otherwise the gaps between groups remained unchanged or even narrowed over 6 years.

Phillips et al. (2002) focused on reading achievement of 187 students in eastern Canada who were followed from Grade 1 (age 6) through Grade 6 (age 12). Students were categorized as below average, average, or above average on the basis of their scores in the first grade. Their findings show that 30% of the students changed category, which challenges the notion that the rich get richer and the poor get poorer. They also report that the initial disadvantage of boys had almost disappeared in the sixth grade.

The study by Luyten et al. (2003) covers a much larger data set than the previous ones (65,129 students) and focuses on the development of general language skills, mathematics, and nonverbal intelligence over a 4-year period (Grades 4 to 8). This study was conducted in the Netherlands and focuses explicitly on the development of children from disadvantaged socioeconomic backgrounds. As mentioned above, this study produces both outcomes that support the notion of a Matthew effect and findings that point to the presence of bigger fish effects. The Matthew effects relate to the increasing gap between students with poorly and well educated parents. On the other hand, the ethnic minority students (Turkish and Moroccan) made up some of their initial disadvantage in nonverbal intelligence. In addition, it was found that schools with poor starting levels made more progress than other schools.

In sum, we can conclude that a number of thorough studies on the presence of Matthew effects have failed to show clear and unequivocal evidence of its presence. For every finding that seems to confirm the notion of a Matthew effect, one can point to a finding that suggests the opposite. The most predominant finding seems to be that students tend to maintain their relative position over the years, although this does not mean that a student's development is predetermined by his or her position at the very start of the school career. Those who start with a disadvantage often evade the lower percentiles at the higher grades. It should also be noted that more in-depth research that has focused on the learning processes underlying the development of reading skills does provide support for the mechanisms described by Stanovich (e.g., Bast & Reitsma, 1997, 1998; Cain, Oakhill, & Lemmon, 2004).

Research Questions

The present study focuses on certain aspects of the presumed Matthew effect that have received relatively little attention in empirical research. First of all, the analyses conducted make a distinction between the school and student levels. It is conceivable that a Matthew effect at the individual level is obscured by an opposing trend at the school level. This may be the case if schools with large percentages of disadvantaged students (which receive extra funding in the Netherlands) manage to make more progress than other schools while at the same time the disadvantage of the initially poor performers within schools continues to grow.

Furthermore, we pay explicit attention to specific groups of students who have been shown to differ with regard to (language) achievement in previous studies. First of all, we explore the impact of socioeconomic background on development in language skills. Numerous studies have shown a substantial correlation between this variable and student achievement, which appears to occur all over the world (OECD, 2004). The question to be addressed here is whether the disparities across socioeconomic lines increase over time. Besides individual background, the composition of the school population is also addressed. According to Stanovich (1986), this factor may increase the growth of initial differences in performance as well. Gender is another variable associated with student performance. Recent studies show a near universal trend across countries of girls outperforming boys with regard to language and reading skills (e.g., Mullis, Martin, Kennedy, & Foy, 2007). However, Phillips et al. (2002) report a strong decrease of the initial boys' disadvantage during the primary school career. Finally, we pay attention to the growth of students who have (implicitly) been identified as either students with learning problems (delayed students) or students with exceptional abilities (accelerated students). The research focuses on the development of general language skills. More specifically, our research questions are the following:

1. To what extent does the initial disadvantage of schools with poor average performance levels increase or decrease over time?
2. To what extent do the initial differences between students within schools increase or decrease over time?
3. How does the achievement gap between students from different socioeconomic backgrounds develop over time?
4. How does the difference in achievement between male and female students develop over time?
5. To what extent does the development of students with delayed and accelerated school careers deviate from the ones with standard careers?

Data Set, Variables, and Analyses

Data Set

For the analyses we make use of the Dutch PRIMA data that were collected in February of 1995, 1997, 1999, and 2001. Starting with 1988–1989, every 2 years data have been collected at a few hundred primary schools in the Netherlands. Since 2007 this collection has changed to a 3-year cycle. Initially, the data were collected for an evaluation of the Dutch educational priority policy (Tesser, Mulder, & Van der Werf, 1991). Each year data on student achievement were collected in Grades 4, 6, and 8. In 1994–1995 the aim of these surveys was broadened to a general monitoring of Dutch primary education and the project was renamed PRIMA. The size of the data collection increased considerably, and pupils from Grade 2 were included for the first time. In the Netherlands, kindergarten and primary education are integrated into a single structure that comprises eight grades. Children enter into the first grade at the age of 4. Since the school year 1994–1995, the data collection has taken place in February, which implies that most of the students in Grade 2 are 6 years old at the time of testing. The entire PRIMA database includes both nationally representative samples and additional samples of schools with disadvantaged student populations (i.e., schools where students with poorly educated parents are overrepresented). Our analyses relate exclusively to the representative samples and include only students in the schools that participated in all four waves of data collection from 1995 until 2001.

The design of the PRIMA cohorts allows for the analysis of individual student growth trajectories. Grade 2 students in 1995 are expected to reappear as fourth graders in 1997, as sixth graders in 1999, and as eighth graders in 2001. However, each year a substantial number of schools end their participation in PRIMA, and these are then substituted for by new ones. The representative samples comprise approximately 400 schools each year, but only 149 schools participated each and every year from 1995 through 2001. More detailed analyses of schools leaving and entering the PRIMA cohorts show no signs of systematic dropout (Roeleveld & Van der Veen, 2007).

Besides entire schools dropping out, one is also faced with the fact that individual students leave the cohort. The reasons for this individual dropout are threefold. Some students move to another school, some students repeat a grade, and some are referred to special education. The students who drop out are substituted by the ones who appear for the first time in one of the higher grades. Most likely these students arrive at the scene for similar reasons that caused the others to drop out. Like the early leavers, these are probably students who have moved from another school or are grade repeaters. The only students who are lost are the ones who are referred to special education. This is about 5% of all students in Dutch primary education.

An estimate of the percentage of grade repeaters among the students who appear on the scene in one of the higher grades can be made on the basis of their birth dates. The students with standard school careers are the ones born between October 1, 1988, and September 30, 1989. A large percentage (44.8%) of the students who appear for the first time in the higher grades are older, which indicates a delayed school career. The percentage is much lower among the other students (viz., 9.4%). Although information on the reasons for students dropping out is not available, it seems plausible that about half of the early leavers moved to another school. The difference in performance level between these students and their nonleaving counterparts is probably quite limited. Most likely, the differences in achievement levels between the early leavers and the students with standard school careers are largely the result of the lower performance of the grade repeaters and the students referred to special education.

In this article we focus on the growth trajectories for language skills of 5,150 students from 149 schools. To be included in the analyses a valid score on at least one point of measurement was required. In addition, students whose date of birth or gender was unknown (n = 112) were excluded. Only 41.3% of all 5,150 students who were included in the analyses took part in the PRIMA data collection at all four points in time. Table 1 provides an overview of the participation patterns and shows that a remarkably large percentage (16.6%) of the students dropped out after the first measurement. This is in line with our knowledge of delayed school careers in the Netherlands. Grade repeating usually occurs in the lower grades and especially after the first two kindergarten years (Reezigt & Knuver, 1995; Roeleveld & Van der Veen, 2007).

Table 1. Student Participation Patterns

	Pattern	*n*	%
1	Present in 1995, 1997, 1999 and 2001	2,129	41.3
	Early leavers		
2	Present only in 1995, 1997, and 1999	274	5.3
3	Present only in 1995 and 1997	379	7.4
4	Present only in 1995	854	16.6
	Late arrivers		
5	Present only in 1997, 1999, and 2001	501	9.7
6	Present only in 1999 and 2001	391	7.6
7	Present only in 2001	295	5.7
	Both late arrivers and early leavers		
8	Present only in 1997 and 1999	107	2.1
9	Present only in 1997	146	2.8
10	Present only in 1999	74	1.4
	Total early leavers (2, 3, 4, 8, 9, 10)	1,834	35.6
	Total late arrivers (5, 6, 7, 8, 9, 10)	1,514	29.4

Variables

The general language skills of the students represent the dependent variable in this study. The language tests taken in each grade have been equated according to item response theory so that student scores in different grades can be related to a common scale. This allows for the estimation of individual growth trajectories.

The tests administered in Grades 4, 6, and 8 comprise three types of items. The first group of items relates to morphology. Students need to indicate whether or not words are structured correctly. One word in a sentence is underlined, and the student is asked to indicate whether or not it is correct use of the Dutch language. The second group focuses on syntax. Students must indicate whether or not a sentence is grammatically correct. The third group is about semantics and requires students to specify if a word or group of words is used in a meaningful context. The sentences and words used in the assignments derive from children's books, magazines, and newspaper sections that are aimed at the appropriate age groups (Mulder, Vierke, & Petersen, 1997).

The language test for the youngest kids (age 6, Grade 2) focuses on the meaning of concepts that are used frequently when learning to read, for example, "many," "few," "first," and "last." The test consists of two parts. The assignments in the first part include all pictures. One example shows four pictures of kids with freckles. Students are asked to mark the picture that shows the kid with most freckles. Another example shows pictures of kids next to each other, kids behind each other, and kids in front of each other. The assignment is to mark the picture with the kids next to each other. The second part focuses on words and sentences. One question shows three words underneath each other. The children are asked to underline the word at the bottom. Another question shows a sentence with one word underlined. The assignment is to underline the next word. The tests in Grade 2 were taken in small groups (no more than 12 children) under the supervision of a trained test leader and in the presence of the children's teacher. The tests were administered during a single morning, in two parts, with a long break in between (Driessen & Claassen, 1996). See Table 2 for the basic statistics with regard to the dependent and independent variables in the analyses.

Mathematics skills were assessed in PRIMA as well, but the present study relates to only language. After the school year 1996–1997 the original mathematics test was replaced. According to Koopman (2002, quoted in Guldemond & Bosker, 2005), the attempts to relate the original and new tests to a common scale were not considered successful.

The independent variables in our analyses are time, socioeconomic background of the student, socioeconomic composition of the school population, gender, and school career (delayed, standard, or accelerated). In the analyses we also take into account whether or not a student has dropped out.

Time is coded in such a way that its effects as estimated in the analyses express the progress made per year. Language scores measured in Grade 2 are denoted by a zero score on the time variable. Scores from Grade 4, which are measured 2 years later, are denoted by a score of 2, and scores from Grades 6 and 8 are denoted by scores of 4 and 6, respectively.

With regard to socioeconomic background, the students are grouped into four categories, primarily based on the educational qualifications attained by their parents. The educational level of the parents is considered low if neither parent attained a qualification above the lower vocational level. Within this group, a distinction is made between students of Turkish or Moroccan descent on one hand and the remaining (predominantly Dutch) students on the other hand. A student is considered Turkish or Moroccan if at least one parent was born in either Turkey or Morocco. If at least one parent has attained a tertiary education diploma, a student is assigned to the upper category. The remaining students end up in the middle category, which is also the largest. See Table 2 for numerical details on the frequency distribution for this variable. Note that only for the students with poorly educated parents is a distinction made between the Turkish or Moroccan students and the others. For 5.2% of the students in the data set, information on socioeconomic background is missing. These have been assigned to the middle category. In the analyses, the effect of socioeconomic background is assessed by means of dummy variables that indicate whether or not a student belongs to a particular category. This means, for example, that if a student belongs to the upper category, he or she gets a score equal to 1 on the well educated parents dummy. All other students get a zero score on this variable. The students in

Table 2. Descriptive Statistics

Language Skills	*M*	*SD*	*n*
Measurements in 1995 (Grade 2)	969.98	34.59	3,565
Measurements in 1997 (Grade 4)	1040.70	36.18	3,425
Measurements in 1999 (Grade 6)	1080.75	34.84	3,224
Measurements in 2001 (Grade 8)	1118.52	34.60	3,188
Independent Variable	**%**		***n***
Socioeconomic background			
Parents' education low; Turkish or Moroccan	8.2		421
Parents' education low; others	32.7		1,684
Parents' education medium	37.6		1,991
Parents' education high	20.5		1,054
School population			
75% or more minority with parents' education low	5.2		269
50%–75% minority with parents' education low	4.7		244
Less than 50% nondisadvantaged; 25%–50% minority with parents' education low	9.2		472
Less than 50% nondisadvantaged; less than 25% minority with parents' education low	33.5		1,723
50%–75% nondisadvantaged	31.8		1,640
75% or more nondisadvantaged	15.6		802
Gender			
Male students	52.7		2,713
Female students	47.3		2,437
School careers			
Students with a delayed school career	19.7		1,019
Students with a standard school career	77.2		3,978
Students with an accelerated school career	3.0		153
Early leavers			
Students leaving the cohort	35.6		1,834

the middle category make up the comparison group. This category is not denoted through a dummy, but the students in this category are distinguished from the rest because they get a zero score on all socioeconomic dummy variables.

The school composition measure is based on the percentages of three categories of students in a school (as counted in 1994). In the Netherlands, schools receive extra funding depending on the number of disadvantaged students. Two categories of disadvantaged students are distinguished: Dutch students with poorly educated parents (neither parent has attained a qualification above the lower vocational level) and minority students with poorly educated parents. The remaining students are assigned to the nondisadvantaged category. In this study schools are grouped into six categories. A zero score is assigned to schools with more than 75% of students in the nondisadvantaged category. In the most disadvantaged schools (Category 5), more than 75% of the students are disadvantaged minorities. With regard to this variable, information is missing for 4.6% of the students. These have been assigned to the median category (less than 50% nondisadvantaged students and less than 25% minority students with poorly educated parents). See Table 2 for more details.

Gender is also measured by means of a dummy variable. Boys get a score equal to 1 on this variable, girls get a 0 score. With regard to school career, three categories are distinguished, namely, delayed, standard, and accelerated. Students are assigned to a particular category on the basis of their age. In the Netherlands, children are assigned to grades primarily by date of birth. Students born before October 1 are usually assigned to a higher grade than the ones who are just a little younger. In the PRIMA data that are analyzed, the students born in the period from October 1, 1998, until September 30, 1989, are the ones with a standard school career. The older students are categorized as delayed and the younger ones as accelerated. In the analyses we use two dummy variables to estimate the growth curves of delayed and accelerated students. The ones with standard school careers make up the comparison group. In most cases the delay is 1 year. The career of 54 students (1%) is delayed by more than 1 year. All accelerated students are only 1 year ahead.

In the analyses we also take into account whether or not a student has dropped out of the cohort. This is done by means of another dummy variable. About half of the early leavers are likely to be grade repeaters, and a small percentage of them are probably referred to special education. It therefore seems likely that the development of the early leavers is less positive than the average. With regard to the early arrivers it is not necessary to include an additional covariate. For this group it is easy to determine whether or not they are grade repeaters because their age is known. This is not the case for dropouts because their school careers follow the standard pattern until they leave the cohort.

Analyses

To answer our research questions, three multilevel models are fitted using the MLwiN software (Rasbash et al., 2000). Three levels are specified in these models. The schools represent the highest level. The students, which are nested within schools, represent the second level. The lowest level is represented by the measurements (maximum of four per student). In the first model, only time is used as an explanatory variable. Thus, it is estimated how much progress students make per year. As the growth in language skills may follow a curvilinear pattern, a quadratic effect of time is estimated in addition to the linear effect. The effect of time is allowed to vary at both the school and student levels, as it seems likely that growth trajectories vary between both schools and individuals.

In the second model, all the covariates mentioned previously are included in the model. In addition, the interaction effects of time with these variables are estimated. The model is extended further by including the interaction effects of the quadratic time variable with the covariates. The main effect of a covariate expresses the (dis)advantage associated with this variable at the first measurement (i.e., in February 1995). The interaction effects of time with the covariates denote to what extent the linear or quadratic growth varies for certain groups of students.

Finally we fit a parsimonious model after removing the nonsignificant effects from the second model. Some restrictions are applied, however. An insignificant main effect is not removed if its interaction with time is statistically significant. An insignificant interaction of a variable with the linear time effect is not removed if the interaction with the quadratic time is significant. An effect is considered statistically significant if it is at least 1.96 times as large as its standard error ($\alpha < .05$ in a two-tailed test).

Results

The findings of the first multilevel analyses are summarized in Table 3. The fixed effects reflect the average growth pattern across all schools and students, whereas the random effects reflect the variation across schools and students. The fixed intercept (grand mean) provides an estimate of the average score in Grade 2 (i.e., when the value of the time variable equals 0). The analyses further yield significant linear and quadratic effects of time. The positive sign of the linear term and the negative quadratic effect suggest a pattern of declining growth, which is in line with the figures presented in Table 1.

The random effects are the most relevant with regard to our research questions as they reveal to what extent the growth patterns differ across schools and students. The variances of the time effects indicate to what extent growth in language skills differs between schools and students.

Table 3. Growth in Language Skills Over Time

Fixed Effects	Effect Estimate	*SE*
Intercept (grand mean)	969.730	1.218
Time—Linear effect	36.749	0.761
Time—Quadratic effect	−2.121	0.114
Random Effects		
School-level variances		
Intercept	170.119	25.342
Time—Linear effect	63.250	9.865
Time—Quadratic effect	1.376	0.222
School-level covariances		
Intercept—Time linear	−35.325	11.987
Intercept—Time quadratic	3.797	1.744
Time linear—Time quadratic	−9.158	1.461
Student-level variances		
Intercept	436.665	23.074
Time—Linear effect	4.607	1.229
Student-level covariance		
Intercept—Time linear	−1.770	4.368
Measurement-level variance		
Intercept	596.272	11.367

With regard to our research questions, the covariances are particularly relevant as they indicate to what extent a high starting level correlates with the rate of improvement. The school- and student-level intercept variances indicate to what extent the starting level of the language scores varies between and within schools. The variance at the lowest level denotes the remaining variation. The school- and student-level variances indicate considerable difference in language scores both between and within schools at the first measurement. The school-level variance (170.119) corresponds with a standard deviation equal to 13.04. The student-level variance corresponds with a standard deviation equal to 20.90. In a standard normal distribution about 30% of the observations (15% on either side) lie outside the range of one standard deviation from the mean score. The score that equals the mean plus one standard deviation thus approximates the median value of the top 30% of the distribution. The value one standard deviation below the mean represents the bottom 30%. With regard to the schools, this implies a difference of more than 26 points (twice the standard deviation) between the top and bottom 30% for the language scores in Grade 2. For students within schools, this difference is close to 42 points. These are substantial gaps, as students gain about 71 points in the first 2 years after the first measurement (see Table 2).

The analyses show significant variance at both the student and school levels of the linear time effect. The quadratic effect turns out to vary only at the school level. Fitting a model with random quadratic effects of time at the student level yields a zero estimate. The random effects at the school level point to a remarkable pattern, although not one that is consistent with a Matthew effect. The negative covariance between the intercept and the linear time effect indicates that linear growth is stronger in schools with a relatively low starting level. The correlation coefficient (r) between the linear effect and the intercept equals –.34 (see Note 1). The quadratic effect is positively correlated with the intercept ($r = .25$), which suggests less decline in growth for schools with a high starting level. The findings also reveal an almost perfectly negative correlation ($r = -.98$) between the linear and quadratic effects of time. This implies that if linear growth is strongly positive, the quadratic effect is strongly negative (i.e., strong linear growth also implies a strong decline in growth in the higher grades and vice versa). This remarkable finding has previously been reported for partly the same cohort of students by Guldemond and Bosker (2005).

The variance in growth at the student level is rather modest in comparison to the school level (4.607 vs. 63.250). The covariance between the linear time effect and the intercept at this level is very small (–1.770) and statistically insignificant as it does not even exceed its standard error (4.368). Its modest size is even more apparent when expressed as a correlation (–.04). This moderate variance implies that within schools, students progress at a fairly similar pace. The negligible correlation between the intercept and time effect at the student level indicates that the limited variation in learning gain within schools is largely unrelated to the students' starting levels.

The outcomes of the multilevel analyses allow for an estimation of the effect of time for schools with different starting levels. The main results are listed in Table 4, which shows the expected language scores over the 1995 to 2001 period in schools with low versus high starting levels. A school with an intercept one standard below the national average is considered a school with a low starting level. These represent the bottom 30%, whereas the schools with an intercept one standard deviation above the national average represent the top 30%. Table 4 also shows the differences between both types of schools. Figure 1 presents a graphical display. Details with regard to the computations are provided in the appendix.

Both Table 4 and Figure 1 show that over time schools with a low starting level make up a considerable portion of their initial disadvantage. The difference in Grade 2 (1995) is nearly halved 4 years later, after which the reduction appears to halt. The pattern suggests even a slight increase from Grade 6 to Grade 8. The largest reduction is realized in the early grades of primary education (from Grade 2 to Grade 4).

Tables 5a and 5b present the results of the analyses that include the effects of the previously mentioned covariates and their interactions with the time variables. The fixed

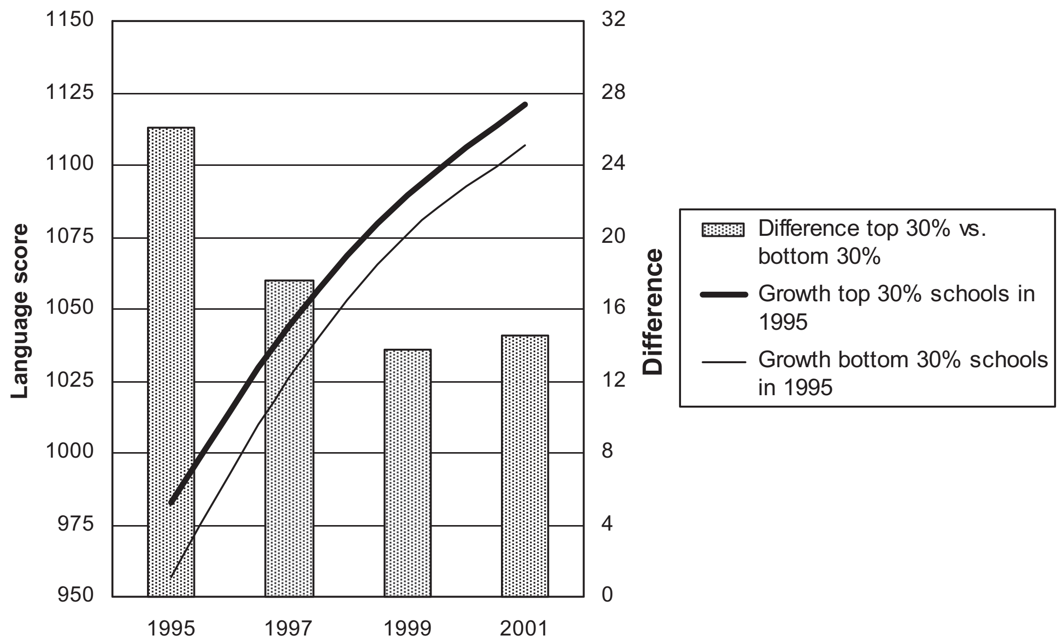

Figure 1. Language growth differences between schools with high and low starting levels

Table 4. Growth in Schools With High and Low Starting Levels

Expected Language Scores	1995	1997	1999	2001
Top 30% schools in 1995	982.77	1043.54	1089.66	1121.14
Bottom 30% of schools in 1995	956.69	1025.95	1075.92	1106.60
Difference top – bottom	26.09	17.58	13.73	14.54

effects are reported in Table 5a. Table 5b relates to the random effects. Two models are fitted. The full model includes the effects of all covariates, whereas in the parsimonious model only the effects significant for $\alpha < .05$ (two-tailed) are retained.

The random effects in Table 5b that are related to both time variables are largely similar to the random effect reported in Table 3. The variances of the school- and student-level intercepts are substantially reduced, however. This indicates that the covariates account for a considerable part of the school- and student-level variance in language scores. At the school level the variance is reduced by 47%, and at the student level this percentage is 23%. The percentages are almost identical for the full model and the parsimonious model.

Nearly all the main effects in Table 5a are significant at .05 level, which indicates that already in Grade 2 (at the age of 6) these variables account for a considerable part of the variation in language skills. The grand mean in the table denotes the language score if the value on all explanatory variables equals 0. More specifically this means the score in Grade 2 for female students with a standard school career who did not leave the cohort, whose parents' education is medium, and who attend a school with at least 75% nondisadvantaged students. The main effects of time also relate to this category of students.

The only variables that do not reveal significant main effects are the ones that relate to the students' school careers. Neither delayed nor accelerated students score significantly above or below the level of the ones with standard school careers. Note that the interaction effects of these variables with time are significant. For the delayed students this indicates growth at a relatively low pace, and for the accelerated students it points to a faster learning rate. The positive quadratic effect associated with delayed school careers implies that the decline in growth is also somewhat less for delayed students. The most plausible explanation for the finding that in Grade 2 the scores of students with delayed and accelerated careers hardly differ from those with standard careers lies in the fact that for most of the delayed students this is a relatively easy year, whereas it probably is a challenging year for the accelerated students. Grade repeating hardly ever occurs before Grade 2. The bulk of the grade repeaters must therefore be doing Grade 2 for the second time in a row, which may explain why they perform at a similar level compared to the other students. The opposite is true for the accelerated students. These students have most likely

Table 5a. Effects of Independent Variables on Language Growth (fixed part)

	Full Model		Parsimonious Model	
Fixed Effects	Effect	*SE*	Effect	*SE*
Intercept (Grand mean)	987.187	1.850	987.456	1.523
Time—Linear effect	36.592	1.433	35.101	0.820
Time—Quadratic effect	−2.227	0.218	−1.959	0.122
Main effects of the covariates				
Low education parents (Turkish or Moroccan)	−26.185	2.325	−24.940	1.618
Low education parents (Others)	−6.587	1.235	−6.228	1.100
High education parents	4.672	1.411	4.432	1.272
Delayed school career	2.008	1.665	1.899	1.659
Accelerated school career	3.895	3.305	4.188	2.937
Early leaver	−16.333	1.052	−16.338	1.050
Disadvantaged school population (6 cat.)	−2.246	0.818	−2.461	0.540
Male student	−7.018	0.999	−7.146	0.907
Interactions with time (linear)				
Low education parents (Turkish or Moroccan)	−0.061	1.670	—	—
Low education parents (Others)	−0.627	0.873	−0.906	0.251
High education parents	0.294	0.982	0.975	0.289
Delayed school career	−5.569	1.079	−5.565	1.064
Accelerated school career	1.980	2.325	1.496	0.707
Early leaver	8.922	1.179	8.927	1.117
Disadvantaged school population (6 cat.)	−0.799	0.652	—	—
Male student	0.957	0.699	1.177	0.213
Interactions with time (quadratic)				
Low education parents (Turkish or Moroccan)	0.118	0.264	—	—
Low education parents (Others)	−0.030	0.139	—	—
High education parents	0.129	0.155	—	—
Delayed school career	0.389	0.161	0.397	0.158
Accelerated school career	−0.084	0.363	—	—
Early leaver	−2.110	0.306	−2.112	0.305
Disadvantaged school population (6 cat.)	0.135	0.099	—	—
Male student	0.038	0.111	—	—

Table 5b

	Full Model		Parsimonious Model	
Random Effects	Effect	*SE*	Effect	*SE*
School-level variances				
Intercept	91.059	15.419	90.991	15.412
Time—Linear effect	64.466	9.953	64.933	9.999
Time—Quadratic effect	1.412	0.225	1.438	0.228
School-level covariances				
Intercept—Time linear	−44.247	10.100	−44.040	10.107
Intercept—Time quadratic	5.290	1.447	5.274	1.454
Time linear—Time quadratic	−9.380	1.479	−9.500	1.492
Student-level variances				
Intercept	336.407	20.838	336.941	20.839
Time—linear effect	2.558	1.161	2.548	1.160
Student-level covariance				
Intercept—Time linear	0.806	4.069	0.825	4.067
Measurement level variance				
Intercept	591.202	11.230	590.788	11.216

skipped the previous grade and now need to catch up to their older classmates.

Most interactions of the covariates with the linear time effect are significant as well. This implies that the linear growth patterns vary across the groups of students denoted by these variables. The only exceptions in this respect are the Turkish and Moroccan students with poorly educated parents. Their considerable disadvantage in Grade 2 (nearly 25 points in the parsimonious model) neither increases nor decreases over time in comparison to the baseline students. The same goes for the school population. This variable points to substantial arrears for students in schools with a disadvantaged majority. The difference between schools with the most and least disadvantaged populations amounts to more than 12 points (5 × 2.461) in Grade 2 and remains stable in the subsequent years.

The interaction of the linear time effect with the education level of the parents indicates a widening gap between the students with well educated versus poorly educated parents. The positive interaction effect implies more progress for the students with well educated parents and the negative interaction effect implies less progress for the ones with poorly educated parents. A similar pattern is found for students with delayed and accelerated school careers. The positive interaction of time with gender indicates that male students compensate for their initial disadvantage in the later stages of the primary school career.

The main effects and the interaction effects for early leavers suggest a complex pattern. They imply that early leavers score more than 16 points below the level of the baseline group in Grade 2. In Grade 4 their disadvantage is less than 7 points, and in Grade 6 it increases again to more than 14 points. Table 6 presents the expected scores per grade for different categories of students. The scores are based on the effects as estimated in the parsimonious model. The development of language scores by socioeconomic categories, school career, and gender are graphically displayed in the Figures 2, 3, and 4.

Table 6 and Figure 2 show that the advantage of students with well educated parents over the comparison group increases over the years. At the same the disadvantage of students with poorly educated parents (except for the Turkish and Moroccan minorities) grows as well. In Grade 2 the difference between both groups is less than 11 points, but in Grade 8 the gap is twice as large. The difference between the Turkish and Moroccan students with poorly educated parents and the other students with poorly educated parents decreases by about 30%, from nearly 19 points in Grade 2 to a little more than 13 points in Grade 8. The disadvantage of the Turkish and Moroccan students remains the same, whereas the other students with poorly educated parents lose ground.

Figure 3 shows the development of students with delayed, accelerated, and standard school careers. Only the measurements in Grades 4, 6, and 8 are taken into account because the situation in Grade 2 is likely to present a biased picture. The large majority of the repeaters in Grade 2 are actually in this grade for the second time. In Grades 4, 6, and 8, only a small minority of the delayed students are in this situation. The figure shows that the gap in language skills between delayed and accelerated students clearly increases over the 4-year period. The difference is nearly 15 points in Grade 4 and has increased to more than 30 points in Grade 8.

The development of language skills by gender shows a closing gap (see Figure 4). In Grade 2 there is moderate disadvantage of 7 points for boys. By the end of primary school the initial disadvantage has disappeared completely.

Table 6. Expected Scores per Student Category (based on the parsimonious model)

	1995	1997	1999	2001
Comparison group	987.19	1049.56	1096.25	1127.27
Low education parents (Turkish or Moroccan)	962.25	1024.62	1071.31	1102.33
Low education parents (Others)	980.96	1041.52	1086.40	1115.61
High education parents	991.62	1055.94	1104.58	1137.55
Accelerated school career	991.64	1057.00	1106.69	1140.70
Delayed school career	989.36	1042.18	1082.51	1110.34
Early leavers	971.12	1042.89	1082.09	—
Students in schools with disadvantaged populations	974.88	1037.25	1083.94	1114.97
Male students	980.04	1044.76	1093.81	1127.19

Conclusion and Discussion

The Matthew effect is a popular concept among educational scholars. It is often referred to as if it were a well-established phenomenon in educational research. This would imply that students who start their school career with a disadvantage are bound to lose ever more ground. The picture presented in thorough empirical studies, however, is more complicated and also less depressing. Stanovich (1986) has put forward compelling arguments that predict ever-increasing arrears of initially disadvantaged students. His line of reasoning largely derives from knowledge on the acquisition of reading skills by new readers. The processes he describes have been confirmed to a considerable extent in subsequent empirical research (e.g., Bast & Reitsma, 1997, 1998; Cain et al., 2004). Still, research aimed at establishing the extent to which achievement gaps

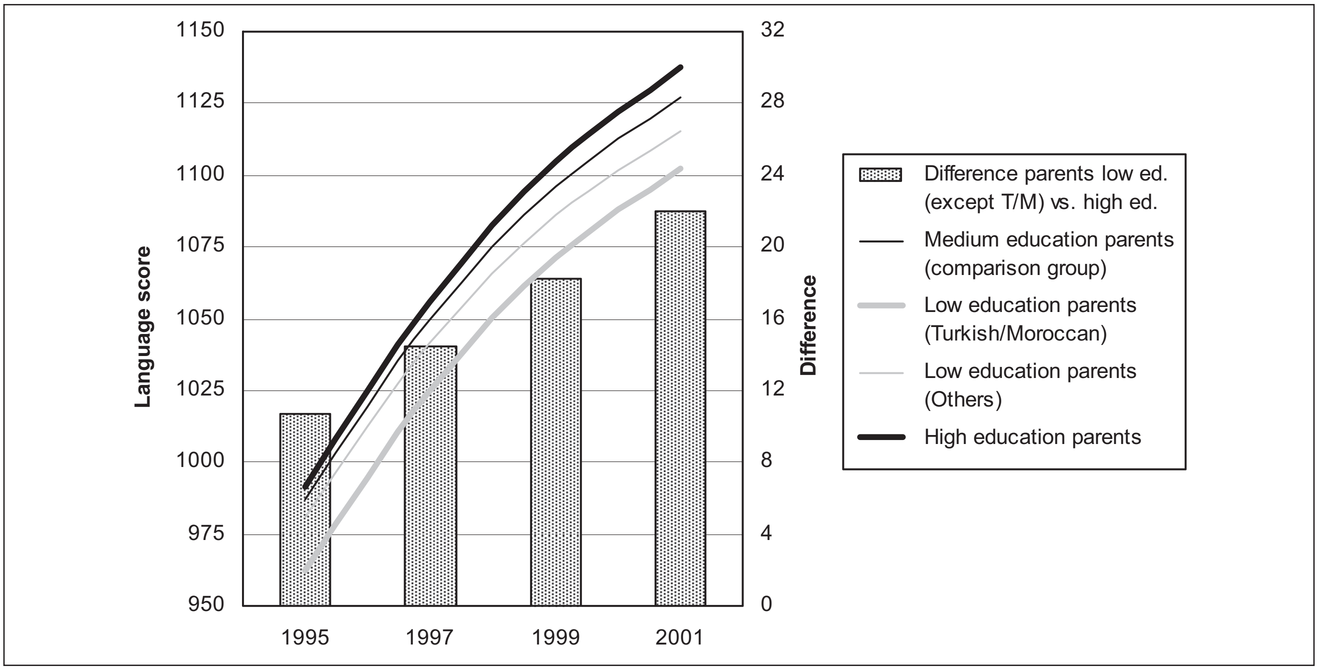

Figure 2. Language growth differences by socioeconomic category

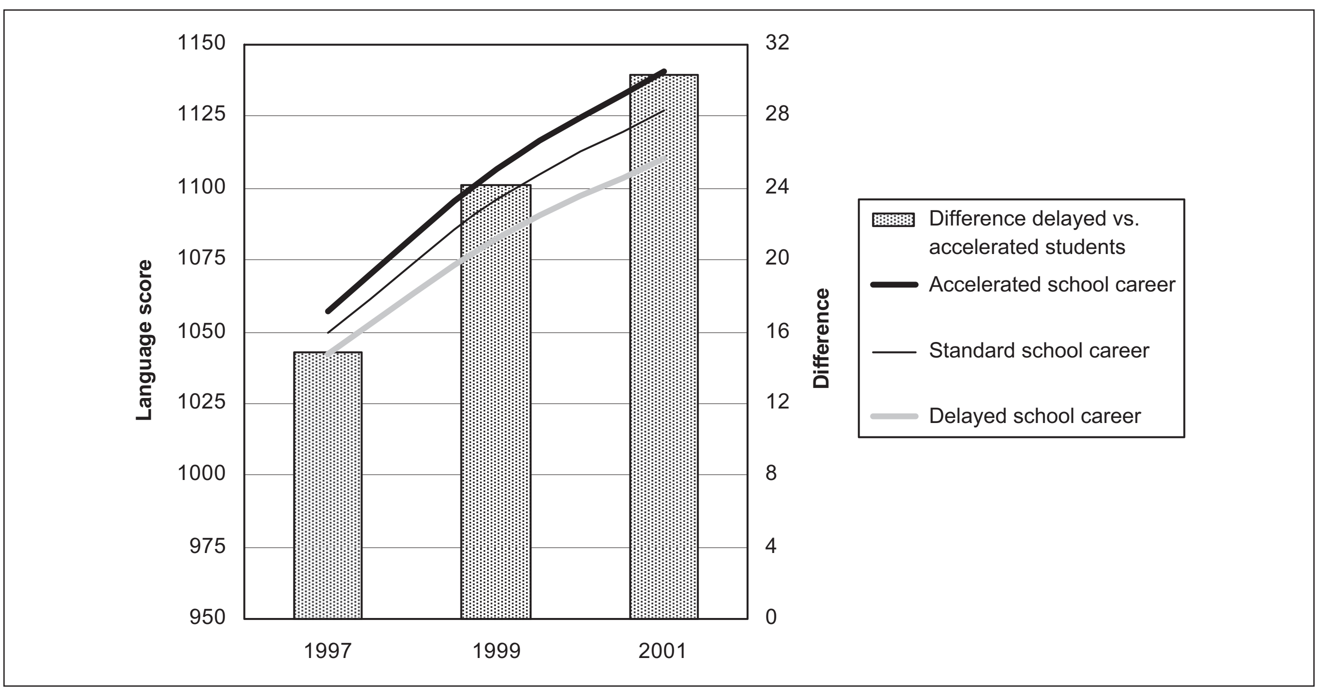

Figure 3. Language growth differences by school career

between low- and high-performing students increase over time has often produced results that contradict the notion that the rich inevitably get richer and the poor poorer. The reduction of initial disadvantages may be referred to as a bigger fish effect. This points to the plain fact that a low starting point provides much room for improvement, whereas decline may be difficult to avoid for the ones starting from a top level.

The outcomes of the present study are clearly in line with the general picture painted in prior research, as they show

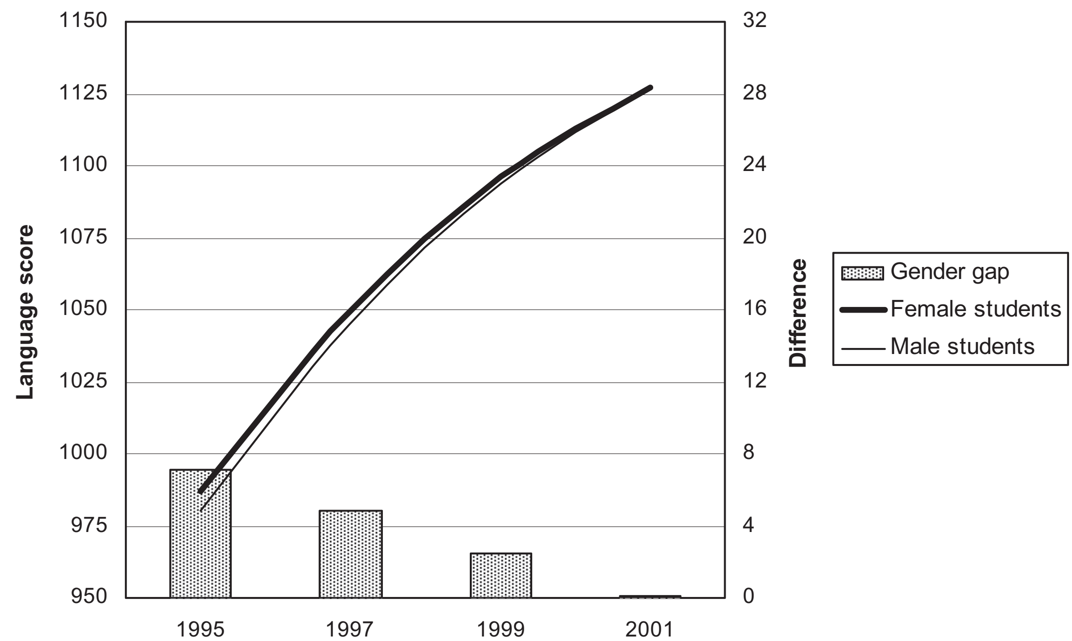

Figure 4. Language growth differences by gender

examples of both Matthew effects and bigger fish effects. The initial disadvantage of schools with poor average performance levels has been found to decrease rather than increase over time. Within schools, the growth patterns of individual students show some variation, but the rate of progress appears to be unrelated to the students' starting levels. Our analyses did find evidence for increases in the achievement gap between students from different socioeconomic backgrounds, although this appears not to apply to the most strongly disadvantaged minorities in the Netherlands (Turkish and Moroccan students). The performance difference between delayed and accelerated students was also found to increase over time, but the initial (and moderate) gender gap turned out to disappear altogether before the end of primary education.

The good news is that initial disadvantages do not necessarily increase over time and may even decline. The challenge for future research is to identify the factors that cause the initial differences to increase or decline. It was beyond the scope of this study to pinpoint such variables, but in the conclusion of this article we suggest some possibilities.

First of all, it seems plausible that the presence of a Matthew effect depends on the specific skills that a study focuses on. It would probably not apply to skills and knowledge that relate to a limited domain (e.g., reading speed or decoding skills). Within a limited domain, students may reach near perfect mastery. This would imply that growth is hardly possible for the ones performing at the highest levels. Progress is achievable only for the poor performers. Further development for the high performers can be achieved only by turning to other domains. This may explain why some studies found a Matthew effect for IQ (arguably the broadest cognitive ability of all) but not for general reading skills (Scarborough & Parker, 2003; Shaywitz et al., 1995) and also why other studies that focused on more specific reading skills found little evidence for a Matthew effect (Aarnoutse & Van Leeuwe, 2000; Bast, 1995).

A basic premise underlying Stanovich's argument is the plausible assumption that knowledge facilitates further learning. In principle this implies that small differences will expand over time. The fact that Matthew effects are often not observed and some studies produce even contradictory findings (bigger fish effects) suggests that counteracting factors are at work as well. Schooling itself may be the main factor in this respect. Research on seasonality of learning consistently shows that cognitive achievement gaps across socioeconomic lines mainly develop during the summer vacation, when school is not in session (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Heyns, 1978). Downey, Von Hippel, and Broh (2004) addresses the effect of schooling on inequality between students from similar socioeconomic backgrounds. They report that also in this respect differences develop mainly during the summer months. The notion that knowledge fosters learning may be valid at a general level, but in school settings this process may be strongly mitigated by teachers who pay special attention to students at risk. This might also account for the closing gender gap that was observed by Philips et al. (2002) and in the present study. It should be noted that although seasonality of learning research provides compelling evidence for the notion that schooling serves as an equalizing force, it does not indicate exactly which aspects of the education process account for its equalizing effects.

Another important assumption behind the presumed Matthew effect relates to the idea that teachers adapt their instruction goals to the performance level of the students. A straightforward application of this principle would indeed imply more challenging goals for high performing students. One would expect that this leads to widening achievement gaps, but if teachers make special efforts to keep the weaker students on track, this process will be weakened to a considerable extent. Still, the present study shows an increasing gap between the accelerated and delayed students. These are students who have been (implicitly) identified as either very fast or rather slow learners. In this respect, our findings are in line with the widening of differences between students with and without learning problems as reported by Scarborough and Parker (2002) and with the findings reported in the contributions to this special issue by Morgan, Farkas, and Wu and by McNamara, Scissons, and Gutknecth. It may be the case that keeping student with severe learning disabilities on track would require more attention and effort than teachers can provide under the circumstances in which they have to do their job. Accelerated students may seek and demand extra intellectual challenges (in and out of school), which would account for their increased learning rates.

Arguably the most important and certainly a reassuring conclusion of educational research on Matthew effects is that they are not inevitable. The challenge for future studies is to indentify factors that are able to prevent the increase of initial disadvantages over time. The focus of these efforts should not be limited to strictly educational variables such as school and classroom practices. Out-of-school factors are likely to play a significant role in the development of achievement gaps as well. The contribution by Forman et al. to this special issue is just one of the numerous studies to support this view.

Appendix

Estimating Time Effects for Schools With Low and High Intercepts

Since the standard deviations of the school intercepts and the linear and quadratic time effect are known, as are their intercorrelations, regression coefficients that specify the growth patterns for a certain value of the school intercept (i.e., the level in Grade 2) can be calculated. Such a regression coefficient (b) can be calculated using the well-known formula

$$b = \text{correlation} \times \text{standard deviation}_{\text{time effect}} / \text{standard deviation}_{\text{intercept}}$$

The above formula can be used to estimate a regression coefficient that indicates how much the linear effect of time deviates from the average effect (i.e., the fixed effect) the more the school-level intercept deviates from the grand mean (i.e., the fixed intercept). In the same way the formula can be used for estimating the quadratic effect of time dependent on the value of the school-level intercept. Table A1 presents the data (school-level variances and covariances) needed to calculate the desired regression coefficients. These figures are derived from Table 3.

The regression coefficients for the linear and quadratic effects thus equal

$$b_{\text{linear}} = -.341 \times 7.953/13.043 = -.208$$
$$b_{\text{quadratic}} = .248 \times 1.173/13.043 = .022$$

In the present case the correlation between linear and quadratic effect of time is almost perfectly negative, so that a linear effect one standard deviation above the mean implies a quadratic effect one standard deviation below the mean. The expected time effects can be calculated as follows for schools in the top 30% and bottom 30% using the formula

$$\text{time effect} = \text{fixed effect} + b \times (\text{school-level intercept} - \text{grand mean})$$

The estimates for the fixed effects and grand mean have been reported in Table 3. The fixed linear time effect equals 36.749, and the fixed quadratic time effect equals –2.121. The grand mean is 969.730. For a school with a score one standard deviation above average in Grade 2 the intercept equals 982.773 (969.730 + 13.043), and for a school one standard deviation below the national average the intercept is 956.687 (969.730 – 13.043). This implies that for a school in the top 30% the difference between its intercept and the grand mean is +13.043, whereas it is –13.043 for schools in the bottom 30%. The expected linear and quadratic time effects for these schools are listed in Table A2.

In conclusion, we illustrate how the expected scores in Grade 6 (i.e., 4 years after the first measurement) have been obtained for schools that were in the top and bottom 30% in Grade 2:

$$\text{Schools in the top 30\%: } 982.773 + 34.041 \times 4 - 1.830 \times 4^2 = 1089.657$$

$$\text{Schools in the bottom 30\%: } 956.687 + 39.457 \times 4 - 2.412 \times 4^2 = 1075.923$$

Table A1. Variances and Covariances

	Variance	*SD*
Intercept	170.119	13.043
Time (linear)	63.250	7.953
Time (quadratic)	1.376	1.173
	Covariance	Correlation
Intercept—Time (linear)	−35.325	−.341
Intercept—Time (quadratic)	3.797	.248

Table A2. Calculation of Time Effects Given Certain Values of the School-Level Intercept

	Top 30% (intercept = 982.773)	Bottom 30% (intercept = 956.687)
Linear effect	36.749 – .208 × 13.043 = 34.041	36.749 + .208 × (–13.043) = 39.457
Quadratic effect	−2.121 + .022 × 13.043 = –1.830	−2.121 – .022 × (–13.043) = –2.412

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note

1. This value is obtained as follows: $r = –35.325/\sqrt{(170.119 \times 63.250)}$.

References

Aarnoutse, C., & Van Leeuwe, J. (2000). Development of poor and better readers during the elementary school. *Educational Research and Evaluation*, *6*, 252–278.

Bast, J. (1995). *The development of individual differences in reading ability* (Unpublished PhD thesis). VU, Paedologisch Instituut, Amsterdam, Netherlands.

Bast, J., & Reitsma, P. (1997). Matthew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research*, *32*, 135–157.

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology*, *34*, 1373–1399.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*, 671–681.

Clay, M. M. (1979). *Reading: The patterning of complex behaviour*. Auckland, New Zealand: Heinemann.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*, 227–268.

Downey, D. B., Von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review*, *69*, 613–635.

Driessen, G., & Claassen, A. (1996). *Voorbereidend lezen, rekenen en schrijven. Toetsresultaten en leerling-en leerkrachtkenmerken in de kleutergroepen van het basisonderwijs* [Preparatory reading, arithmetic and writing. Test results and student and teacher characteristics in the kindergarten grades of primary education]. Nijmegen, Netherlands: ITS.

Guldemond, H., & Bosker, R. J. (2005). *Ontwikkelingen in de groei van achterstanden? Op zoek naar de rol van de school* [Developments in the growth of disadvantages? In search of the school role]. Groningen, Netherlands: GION.

Heyns, B. (1978). *Summer learning and the effects of schooling*. New York, NY: Academic Press.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*, 437–447.

Lundberg, I. (1984). *Learning to read (School research news letter)*. Stockholm, Sweden: National Board of Education.

Luyten, H., Creemers-van Wees, L. M. C. M., & Bosker, R. J. (2003). The Matthew effect in Dutch primary education: Differences between schools, cohorts and pupils. *Research Papers in Education*, *18*, 167–195.

Merton, R. K. (1968). The Matthew effect in science. *Science*, *159*, 56–63.

Mulder, L., Vierke, H., & Petersen, B. (1997). *De vierde fase van de OVB-cohortonderzoeken. Het leerling-, school- en klasonderzoek in schooljaar 1994/1995* [The fourth stage of the OVB cohort surveys. The student, school and class survey in school year 1994/1995]. Nijmegen, Netherlands: ITS/GION.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary schools in 40 countries*. Boston, MA: TIMSS & PIRLS International Study Center.

Organisation for Economic Co-operation and Development. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris, France: Author.

Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first to sixth grades. *Journal of Educational Psychology*, *94*, 3–13.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., . . . & Lewis, T. (2000). *A user's guide to MLwiN* (Multilevel Models Project). London, UK: University of London.

Reezigt, G. J., & Knuver, A. (1995). Zittenblijven in het basisonderwijs [Grade repeating in primary education]. *Pedagogische Studiën*, *72*, 114–132.

Roeleveld, J., & Van der Veen, I. (2007). Kleuterbouwverlenging in Nederland: Omvang kenmerken en effecten [Extended kindergarten in the Netherlands: Size, characteristics and effects]. *Pedagogische Studiën*, *84*, 448–462.

Scarborough, H. S., & Parker, J. D. (2003). Matthew effects in children with learning disabilities: Development of reading, IQ and psychosocial problems from Grade 2 to Grade 8. *Annals of Dyslexia*, *53*, 47–71.

Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., Stuebing, K. K., Francis, D. J., & Shaywitz, S. E. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly*, *30*, 894–906.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Tesser, P., Mulder, L., & Van der Werf, G. (1991). *De eerste fase van de longitudinale OVB-onderzoeken, het leerlingonderzoek* [The first stage of the longitudinal OVB studies, the student survey]. Nijmegen, Netherlands: ITS/RION.

Toonder, M. (1983). *Geld spelt geen rol* [Money does not matter]. Amsterdam, Netherlands: De Bezige Bij.

Walberg, H. J., & Tsai, S. L. (1983). Matthew effects in education. *American Educational Research Journal*, *20*, 359–373.

Wechsler, D. (1989). *Wechsler intelligence scale for children – Revised.* San Antonio, TX: The Psychological Corporation.

Woodcock, R. & Johnson, M.B. (1978). *Woodcock-Johnson Psychoeducational Battery.* Allen Park, TX: DLM Teaching Resources.

About the Authors

Hans Luyten, is an Associate Professor at the Faculty of Behavioral Sciences (Department of Educational Organization and Management) of Twente University in the Netherlands. His research interests include educational effectiveness, social inequality, educational policy and international comparisons. Innovative research methodology is a key characteristic in much of his work.

Gerdy ten Bruggencate, is currently working at the teacher training institute ELAN of Twente University in the Netherlands. Her research interests include educational effectiveness and the impact of school leaders on school performance.