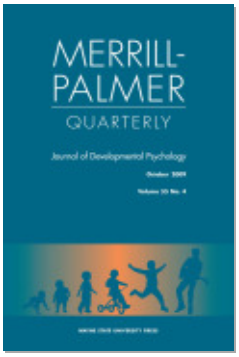
# School Engagement and Language Achievement: A Longitudinal Study of Gender Differences across Secondary School




Eva Van de gaer, Heidi Pustjens, Jan Van Damme, Agnes De Munter

Merrill-Palmer Quarterly, Volume 55, Number 4, October 2009, pp. 373-405 (Article)

Published by Wayne State University Press
*DOI: 10.1353/mpq.0.0034*

# School Engagement and Language Achievement

## A Longitudinal Study of Gender Differences across Secondary School

Eva Van de gaer *K.U.Leuven, Belgium and ACER, Australia*
Heidi Pustjens, Jan Van Damme, and Agnes De Munter *K.U.Leuven, Belgium*

The present study investigated (1) gender differences in the longitudinal development of language achievement and school engagement (i.e., effort for language, attitude toward learning tasks, interest in learning tasks, and relationship with teachers) across secondary school (Grades 7–12, ages 12–18) and (2) gender differences in the association between these developmental processes. Data were drawn from the Longitudinal Project in Secondary School (LOSO project) that followed a cohort of 2,270 students in Flanders (Dutch-speaking part of Belgium) during secondary school. We used univariate and multivariate latent growth curve models to analyze the longitudinal data. The results showed, as expected, significant gender differences in the development of language. Girls showed a quasi-linear positive learning gain in language across secondary school, whereas boys started with a decline followed by acceleration in their learning gain in language. With regard to school engagement we observed a decline both for boys and girls, but the decline in the effort for language and the attitude toward learning tasks was steeper for boys than for girls. In addition, we found evidence for a positive longitudinal association between language achievement and school engagement, meaning that students who showed a smaller decline in school engagement also showed more learning gains in language. Moreover, only for boys were the school engagement starting levels related to their learning rates in language.

Eva Van de gaer, Centre for Educational Effectiveness and Evaluation and Australian Council for Educational Research; Heidi Pustjens and Jan Van Damme, Centre for Educational Effectiveness and Evaluation; Agnes De Munter, Research Centre Women and Education.

Correspondence should be addressed to the first author at Australian Council for Educational Research, 19 Prospect Hill Road, Camberwell VIC 3124, Australia. Phone: (03) 9277-5489. Fax: (03) 9277-5500. E-mail: Vandegaer@acer.edu.au.

*Merrill-Palmer Quarterly*, October 2009, Vol. 55, No. 4, pp. 373–405. 

During the past three decades the focus of educational research on gender differences has shifted from the underachievement of girls to the underachievement of boys (Barnett & Rivers, 2006). Girls now outperform boys on almost every educational outcome (Marks, 2008; Mead, 2006; Stowe, Arnold, & Ortiz, 2000). Especially in language areas the gender differences are large and pervasive, and these gender differences are considered to be a major factor in the origin of gender differences in national exams, the widening of the gender gap in college attendance, and degree attainment rates (Burgess, McConnell, Propper, & Wilson, 2004; Sum, Kirsch, & Taggert, 2002).

One of the crucial determinants of boys' underachievement is their lower levels of school engagement (Arnot, David, & Weiner, 1999; Clark & Trafford, 1995; Davies & Brember, 2001; Engels, Aelterman, Van Petegem, & Schepens, 2004; Lamb, 1997; Walsh, Hickey, & Duffy, 1999). According to Finn (1993), low levels of achievement and, eventually, school dropout may be the end point of a developmental process of disengagement and withdrawal from school that may have begun in the early years of schooling (see also Fredricks, Blumenfeld, & Paris, 2004; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002). To test this hypothesis, we need to investigate the developmental trajectories of school (dis)engagement and those of achievement. However, most studies to date that have investigated these issues have relied on cross-sectional methods, which cannot provide insight into the way educational outcomes, such as school engagement and achievement, change over time.

The present study aims to fill this gap in the literature by tracking the same group of students over time and by investigating the association between changes in school engagement and changes in language achievement. Previous research has focused either on the longitudinal development of language or on the development of school engagement, but very few studies have investigated the link between these two developmental processes. In addition, we focus on gender differences. We want to address the question of whether boys become more disengaged with schooling than girls over time and whether this negative development is accompanied by a decline in boys' learning rates in language. In order to answer these questions, we will provide longitudinal data on the development of language achievement and of school engagement measures (i.e., attitude toward homework, effort for language, relationship with teacher, interest in learning tasks) for boys and girls across secondary school in Flanders (the Dutch-speaking part of Belgium). In Flanders, secondary school consists of six years (equivalent to U.S. Grades 7–12), and students start secondary school at age 12.

### *Gender Differences in the Development of Language and of School Engagement*

Studies investigating gender differences in language achievement at different ages drive home to the point that girls achieve better than boys in language. However, only small to nonexistent gender differences exist during elementary school, whereas the gender differences are larger at the end of secondary school (Cole, 1997; Coley, 2003; Hill & Russell, 1999). Although these results suggest that girls have higher learning rates in language, only longitudinal research, which tracks the language achievement of the same group of students over time, can test such a claim. However, longitudinal evidence of gender differences in learning rates in language is still scarce. One of the few exceptions is Kiplinger's (2004) study, which investigated students' growth in reading and writing achievement from Grade 3 to Grade 10 and found that only at the high school level did girls progress at a significantly faster rate in reading.

Although the research results on the development of gender differences in language achievement are consistent, the research results on gender differences in the development of school engagement are not. Van Schooten, de Glopper, and Stoel (2004) investigated the development of attitudes toward literature across Grades 7–11 in the Netherlands using structural equation models (SEMs). They measured behavioral intentions such as "I want to read many literature works in the future" and enjoyment of reading and found a stronger decline for boys than for girls. In contrast, Watt (2004a) used complex growth curve models to investigate the development of self-reported effort for language and interest in language of Australian students and found no support for greater declines among boys than among girls between Grades 7 and 11. On the contrary, Jacobs et al. (2002) used quadratic growth curve models to investigate the development of language task values such as interest, importance, and usefulness in American students across elementary school and found more complex gender differences. They found a steeper decline in language task values for girls up to the end of elementary school but an increase of girls' task values during high school, whereas boys' task values leveled off.

To conclude, the inconsistencies in the research results on gender differences in the development of school engagement may be due to the variability in methodology (e.g., SEM, growth curve analysis, linear versus quadratic growth curves, etc.) and sample characteristics across existing studies, such as differences in cohorts of students or in the measurement of school engagement. The present study will investigate gender differences in the development of language achievement and four school engagement

measures by tracking the same group of students over time and by using the same methodology. Doing so will make the findings of achievement and the four measures of school engagement comparable and interpretable.

In addition, the present study will focus on the associations between the development of language achievement and the development of school engagement for boys and girls using longitudinal data. Most research linking gender differences in school engagement to gender differences in language achievement is cross-sectional. Prior cross-sectional research has shown that school engagement and behavior such as time spent on homework, attention in the classroom, and interest in learning tasks of girls are higher than those of boys. When these gender differences in school engagement are taken into account, the gender differences in achievement become smaller (Anderman & Maehr, 1994; Downey & Yuan, 2005; Freudenthaler, Spinath, & Neubauer, 2008; Van de gaer, Pustjens, Van Damme, & De Munter, 2007; Stowe et al., 2000; Whitelaw, Milosevic, & Daniels, 2000). In addition, school engagement plays a more important role for boys than for girls. For example, Freudenthaler et al. (2008) found that school-related intrinsic motivation predicted boys' achievement but not that of girls. Likewise, Van de gaer et al. (2007) showed that boys' motivation toward learning tasks and concentration in the class were stronger predictors of boys' achievement. Girls' achievement depended less on school engagement such as interest in learning tasks. Freudenthaler et al. (2008) suggested that girls' lower dependence on school achievement conveyed a substantial advantage for girls, as students need to engage in learning tasks that they do not always like.

Stoel, Peetsma, and Roeleveld (2003) conducted one of the few studies that used longitudinal data to investigate the association between the development of language achievement and the development of school engagement. They found a positive relation between the development of school investment and self-confidence and the development of language achievement in elementary school in the Netherlands. This result indicates that students who showed less of a decline in school investment and self-confidence showed higher learning rates in language. However, the study focused on elementary school, a time period in which gender differences are small to nonexistent. The present study examined the association between change in four indicators of school engagement and change in language achievement for boys and girls during secondary school.

### *Research Hypotheses*

In the present study, we used a longitudinal approach to investigate the development of school engagement and language achievement in second-

ary school (i.e., Grades 7–12) and the association between these two developments. To this end, we will address four research questions.

First, we will examine the development of language and of school engagement separately without looking at gender differences. We will investigate four indicators of school engagement: effort for language, the attitude toward homework, the interest in learning tasks, and the relationship with teachers. Based on previous research (Eccles & Midgley, 1989; Hill & Russell, 1999; Olson, 2002; Watt, 2004a, 2004b; Wigfield & Eccles, 2000; Wigfield et al., 1997), we expect a decline in students' school engagement and an increase in language achievement across secondary school (Hypothesis 1). To investigate the changes in language achievement and in school engagement, we will investigate the nature of the development by modeling data using growth curve analyses. Previous research indicated that linear growth curves are often not complex enough to grasp the gender differences in the development of school engagement (Fredricks & Eccles, 2002; Jacobs et al., 2002; Kolaweski-Jones & Duncan, 1999). Therefore, we will apply a linear growth curve in a first model. We then test a second model that uses a more complex growth curve, namely a curvilinear or quadratic growth curve assuming that the data fit a curvilinear model. Curvilinear models allow for the identification of more complex developmental patterns (as opposed to a linear developmental pattern) such as a deceleration or acceleration in the rate of decline over time. By comparing these two models (linear versus quadratic), we can test which model shows a better fit to the data.

Second, we will investigate gender differences in the development of language achievement and in the development of school engagement. We expect that boys' decline in school engagement will be steeper than that of girls and that girls will make more progress than boys in language achievement (Hypothesis 2).

Third, we will investigate the relationship between the development of school engagement and the development of language achievement across secondary school without looking at gender differences. Using multivariate latent growth curve analysis (Sayer & Willett, 1998), we will model the growth curves of school engagement and language achievement simultaneously. We expect a positive relationship between the starting levels and the rates of change in language achievement and school engagement (Hypothesis 3). In other words, we expect that students starting secondary school with a higher initial school engagement will start secondary school with higher language achievement and that students showing a smaller decline in school engagement will show a greater increase in language achievement.

Finally, we will examine whether boys and girls differ in the association between the development of language and school engagement. Based

on previous research (Freudenthaler et al., 2008; Van de gaer et al., 2007), we expect that the associations between the starting levels and the rates of change of school engagement and language achievement will be stronger for boys than for girls (Hypothesis 4).

## Method

### *Sample and Procedure*

Data were taken from a large-scale study, the Longitudinaal Onderzoek Secundair Onderwijs (LOSO) project (Longitudinal Research Project in Secondary Education) that started in 1990 and followed a cohort of students during secondary school and afterward (age 12–21) in Flanders, the Dutch-speaking part of Belgium (Van Damme & Onghena, 2002). In a first step of LOSO, schools were selected to participate in the project. This selection was based on school characteristics that were representative for Flanders such as the size of the school, the type of school, the curriculum offered, and the representation of both private Catholic schools and public schools. Next, all the students in the selected schools were solicited to participate. Less than 1% of the parents did not return consent forms for their children to participate in the study.

The analyses of the present study are based on a sample of 2,270 students who started Grade 7 in 1990 and reached Grade 12 six years later. To rule out possible confounding factors that may affect change in language achievement and in school engagement, such as changing between tracks and moving to another school, we only considered students who did not move to another school and who stayed in the same track during the studied time period (i.e., Grade 7–12). Because we restricted our sample to students who followed the academic and the technical track,[1] the final sample consisted of more girls ($N = 1{,}327$) than boys ($N = 943$), as more girls than boys choose the academic track. Information on the income and the educational and occupational levels of both parents indicates that the students are from middle-class backgrounds. The mean age of the participants was age 12 in Grade 7 and age 18 in Grade 12. Ninety-five percent of the students spoke only Dutch at home. Boys and girls did not differ significantly in background characteristics. We found no significant gender differences in

[1] In Flanders, students are tracked from Grade 9 into four possible tracks: academic, technical, artistic, and vocational. We decided to include only students from the academic and technical tracks because (a) the LOSO data included too few students from the artistic track and (b) the language curriculum of the vocational track differed substantially from that of the academic and technical tracks, making comparisons very hard.

socioeconomic status ($t[1, 2{,}270] = 1.53$, $p > .05$). When we compared the percentages of boys and girls who started secondary school with delay (boys, 3.08%; girls, 2.56%) versus without delay and who spoke Dutch (native language) versus another language at home (boys, 4.88%; girls, 5.58%), we found no significant differences ($\phi = 0.02$, $p > .05$, and $\phi = -0.02$, $p > .05$, respectively).

### *Measures*

*Language achievement.* During the LOSO project, we tested students' language achievement five times: at the beginning and at the end of Grade 7 and at the end of Grades 8, 10, and 12. As there are no national exams in Flanders, the language achievement tests were especially constructed for the LOSO project. The tests are composed of curriculum relevant multiple-choice items and are approved by a board of inspectors and teachers. We constructed different versions of the language achievement tests within and between grades in order to take curriculum differences within and between grades into account. The majority of the test items tapped spelling and reading comprehension. Using multiple-choice questions, we asked students to indicate which word was spelled correctly and to read a short text and answer content-related questions afterward. Thus, the language achievement tests tapped spelling and reading comprehension but not writing.

To compare the scores on the different versions of language achievement tests across grades, we calibrated the test scores by item response theory (IRT) analyses using BIMAIN (Zimowski, Muraki, Mislevy, & Bock, 1994). This resulted in IRT scores that are measured on the same scale for language ability on each of the five measurement occasions (i.e., beginning of Grade 7, end of Grade 7, and end of Grades 8, 10, and 12). Table 1 shows the Cronbach's alphas of the different language achievement tests that ranged from 0.66 to 0.93, indicating moderate to very good reliabilities.

*School engagement.* At the end of Grade 7, Grade 8, Grade 10, and Grade 12[2] we administered a well-being questionnaire to all students. The

[2] Note that language achievement was measured five times, whereas the engagement variables were only measured four times during secondary school. This causes no problem when analyzing the associations between the growth trajectories of language achievement and engagement because Mplus can accommodate for differing numbers of waves in language achievement and behavioral and emotional engagement (Sayer & Willett, 1998). In addition, growth curve models assume that there is a latent growth trajectory that underlies the observed outcomes. This means that in the present study a growth curve will be estimated for the whole period of secondary school based on four observations (i.e., end of Grades 7, 8, 10, and 12). Thus, also for time points that were not actually observed such as the beginning of Grade 7, Grade 9, and Grade 11 or in the middle of Grade 8, the growth curve produces estimates.

**Table 1.** Descriptive Statistics for the Repeated Measures of Language Achievement, Interest in Learning Tasks, Relationship with Teachers, Attitude toward Homework, and Effort

| | Grade 7 (Start) | Grade 7 (End) | Grade 8 (End) | Grade 10 (End) | Grade 12 (End) |
|---|---|---|---|---|---|
| *Language achievement* | | | | | |
| *M* | –0.07 | –0.10 | –0.14 | 0.02 | 0.30 |
| *SD* | 0.64 | 0.66 | 0.70 | 0.71 | 0.69 |
| Reliability | 0.93 | 0.80–0.90 | 0.82–0.90 | 0.66–0.74 | 0.77–0.83 |
| Skewness | –0.27 | –0.31 | –0.94 | –0.31 | –0.56 |
| Kurtosis | 0.79 | 0.21 | 1.31 | 0.5 | 1.11 |
| *N* students | 2180.00 | 2167.00 | 2150.00 | 2155.00 | 2092.00 |
| *Interest in learning tasks* | | | | | |
| *M* | | 3.66 | 3.44 | 3.06 | 2.99 |
| *SD* | | 0.61 | 0.68 | 0.63 | 0.60 |
| Reliability | | 0.83 | 0.86 | 0.87 | 0.86 |
| Skewness | | –0.17 | –0.33 | –0.14 | –0.20 |
| Kurtosis | | 0.23 | 0.33 | 0.35 | 0.32 |
| *N* students | | 2152.00 | 2157.00 | 2043.00 | 2010.00 |

| | | | | |
|---|---|---|---|---|
| *Relationship with teachers* | | | | |
| *M* | 3.81 | 3.73 | 3.55 | 3.46 |
| *SD* | 0.58 | 0.62 | 0.54 | 0.55 |
| Reliability | 0.85 | 0.88 | 0.88 | 0.87 |
| Skewness | –0.29 | –0.48 | –0.28 | –0.63 |
| Kurtosis | 0.44 | 0.84 | 0.58 | 1.24 |
| *N* students | 2153.00 | 2158.00 | 2043.00 | 2013.00 |
| *Attitude toward homework* | | | | |
| *M* | 4.27 | 4.24 | 3.95 | 3.80 |
| *SD* | 0.62 | 0.65 | 0.67 | 0.70 |
| Reliability | 0.77 | 0.82 | 0.83 | 0.82 |
| Skewness | –0.88 | –1.05 | –0.49 | –0.48 |
| Kurtosis | 0.63 | 1.57 | 0.18 | 0.13 |
| *N* students | 2152.00 | 2154.00 | 2043.00 | 2010.00 |
| *Effort* | | | | |
| *M* | 3.21 | 3.07 | 3.02 | 2.90 |
| *SD* | 0.76 | 0.77 | 0.74 | 0.75 |
| Reliability[a] | | | | |
| Skewness | –0.76 | –0.54 | –0.34 | –0.26 |
| Kurtosis | 0.28 | –0.11 | –0.27 | –0.29 |
| *N* students | 1992.00 | 2077.00 | 1746.00 | 1860.00 |

[a]The reliability of effort is not given since this variable was measured by one item.

well-being questionnaire was an adaptation of the *schoolvragenlijst voortgezet onderwijs* (school questionnaire secondary education) of Smits and Vorst (1982), supplemented with a number of new items (formulated with the help of a questionnaire by Janssen, 1982). Using factor analysis, we constructed three scales composed of the same items at the four measurement occasions. These are the interest in learning tasks (e.g., "I enjoy most of the subjects taught in this school"), relationship with teachers (e.g., "I think that most of the teachers are very helpful when I have problems with school work"), and attitude toward homework (e.g., "When I have homework, I start as soon as possible"). Appendix 1 is a full list of the items of each of the scales. The Cronbach's alphas at the different grades are shown in Table 1.

In addition, we asked teachers who taught language to indicate on a 4-point scale (ranging from a lot of effort to no effort at all) to what extent each student put effort into the subject language. We used these teacher reports as a measure for the effort for language.

Based on the definition of school engagement of Fredricks et al. (2004), we can consider effort for language and attitude toward homework as two indicators of behavioral engagement. In contrast, we consider relationship with teachers and interest in learning tasks as two indicators of emotional engagement. Considering effort as an indicator of behavioral engagement needs some further explanation because it can be considered as an indicator of behavioral as well as cognitive engagement. Fredricks et al. (2004) make a distinction between effort that is primarily behavioral (e.g., doing the work) and effort that is more cognitive and aimed at learning and mastering the material (e.g., to go beyond the requirements, perseverance, psychological investment). In the present study, we did not explicitly ask teachers to indicate to what extent students go beyond the requirements. Therefore, we assumed that the measurement of effort in the present study corresponds more to behavioral engagement than to cognitive engagement.

### *Missing Values*

In Grade 7, 8, 10, and 12 students' language abilities were tested, and questionnaires were administered assessing school engagement. As in most longitudinal studies, the project suffers from attrition. For example, students were sometimes absent on the days that the tests were administered. We included students with missing data in the sample if they had at least one out of four observations on the repeated measures, which severely limited the loss of students due to panel attrition. Two students had no single measure of language achievement; three students had no measure on the vari-

ables "attitudes toward homework," "interest in learning tasks," and "relationship with teachers" at any of the measurement occasions; and "the effort for language" variable was missing at all measurement occasions for only eight students. Although the number of missing data increased by grade (see Table 1), we applied maximum likelihood estimation, which can handle such patterns of missing data at random (Peugh & Enders, 2004). Thus, the maximum available information at each measurement occasion was used to obtain estimates of the model parameters.

### *Data Analytical Strategy*

We used latent growth curve analysis to analyze our longitudinal data (Muthén & Khoo, 1998). This statistical technique allows estimating of individual growth curves that represent change or growth in language achievement and in school engagement across the entire period of secondary school based on four (or five) measurement occasions. With linear growth curves, two growth parameters are estimated, namely an initial level growth parameter (i.e., intercept) and a growth rate parameter (i.e., slope). These parameters are viewed as latent variables, and these latent variables can be estimated using SEM. Both parameters vary between individuals, meaning that for each individual a growth curve is estimated with a specific initial status and a specific rate of change. Because we have (at least) four waves of data, we can also estimate quadratic growth curves that can identify more complex trajectories than linear growth curve models (Singer & Willett, 2003). Quadratic models include one extra growth parameter, known as the quadratic growth curve parameter or the curvature growth curve parameter. The interpretation of the intercept in quadratic growth curve models is the same as in the linear growth curve models, but the interpretation of the slope or growth rate parameter is somewhat different. In the quadratic growth curve models there is no constant common slope because the growth changes smoothly over time. Therefore, the slope can be interpreted as an instantaneous rate of change or growth that can change over time (as indicated by the quadratic growth curve parameter). We centered the intercepts of language achievement scores and of the four school engagement variables at the beginning of Grade 7.

The analyses were conducted in four steps. In a first step we fitted univariate latent growth models, meaning that latent growth curve models are estimated separately for language achievement and for the four school engagement variables (cf. Hypothesis 1). We determined the adequate form of each growth trajectory by comparing the fit between two nested models, namely one specifying a linear form and the other specifying a quadratic

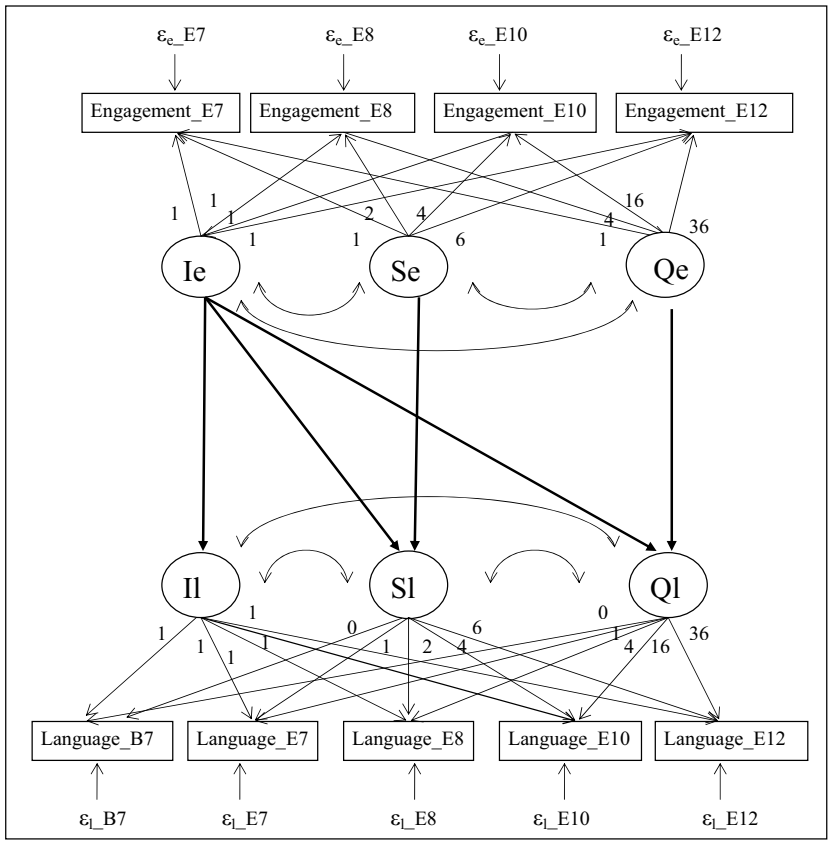
form. The beginning of Grade 7 was chosen as the reference point, meaning that the intercept or the initial status refers to the start of secondary school. The unequal intervals between the different measurement occasions were specified in the model in order to obtain adequate estimations of the growth curve parameters. The univariate latent growth models extrapolate individual growth curves across the entire period of secondary school based on four (or five) measurement occasions.

In a second step we examined whether the growth trajectories in language achievement and in the four school engagement variables differed between boys and girls (cf. Hypothesis 2). To this end, a dummy coded variable "gender" was entered as a predictor of the latent growth curve factors (intercept, linear and quadratic factor).

In a third step we estimated multivariate latent growth curve models (cf. Hypothesis 3). Figure 1 shows the path diagram of the multivariate latent growth model. The variables "attitudes" and "language" represent the repeated measures of school engagement and language. The repeated measures are represented by three latent growth factors, namely an intercept (I), a slope (S), and if applicable a quadratic factor (Q), that were allowed to covary with each other. Multivariate models estimate both developmental processes simultaneously and thus allow us to investigate whether the change in language achievement is associated with the change in school engagement (Singer & Willett, 2003). The growth parameters of language achievement were modeled as the dependent variables, whereas growth parameters of (one) school engagement variable were modeled as the independent variables. In the model depicted in Figure 1, the initial status, the slope, and the quadratic term of the school engagement variable predicted the initial status, the slope, and the quadratic term of language achievement, respectively. The initial status of engagement also predicted the slope and the quadratic term of language achievement.

In a fourth step we used multiple group analysis to investigate whether the associations between the growth parameters in school engagement and the growth parameters in language achievement were similar for boys and girls (cf. Hypothesis 4).

The analyses were carried out by full information maximum likelihood estimation using the Mplus software package version 3.11 (Muthén & Muthén, 1998–2004). Although our data are hierarchically structured because students are nested within schools, we did not model a school level because investigating the effect of school characteristics on change in language achievement and in school engagement is not the focus of the present study. Instead we used an estimation method (i.e., full-information maximum likelihood) that takes into account the hierarchical structure of




*Note*. Ie, Se, Qe, and Il, Sl, Ql stand for the intercept, slope, and quadratic term of school engagement and language achievement, respectively. B7, E7, E8, E10, and E12 stand for beginning grade 7, end Grade 7, end Grade 8, end Grade 10, and end Grade 12. $\varepsilon_l$ and $\varepsilon_e$ stand for error in language achievement and error in school engagement, respectively.

**Figure 1.** Multivariate Latent Growth Curve Model of Language Achievement and Engagement.

the data (or the nonindependence of the data) by appropriately adjusting the standard errors and the chi-square tests of model fit (Muthén & Muthén, 1998–2004). As recommended by Hu and Bentler (1999), we used multiple indices in judging the model fit. Because the conventional overall tests of fit ($\chi^2$ tests of model fit and also the $\chi^2$ difference tests) are very dependent on the sample size, the model fit will be judged in combination with other fit indices such as the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square error of approxima-

tion (RMSEA) that have proven to be less dependent on sample size, robust against model misspecification, or relatively stable under different estimation methods (Fan, Thompson, & Wang, 1999; Hu & Bentler, 1998). We considered a fit greater than 0.90 and 0.95 of CFI and TLI, respectively, as acceptable and excellent fit to the data. RMSEAs less than 0.05 are indicative of a close fit, and values up to 0.08 are considered as reasonable errors of approximation (Hu & Bentler, 1999; Maruyama, 1998). Chi-square difference tests were used to evaluate differences in model fit between two nested models.

## Results

### *Descriptive Statistics*

Table 1 shows the means, standard deviations, reliabilities, skewness, kurtosis, and number of students for each variable examined at each time point. The skewness and the kurtosis[3] of all the variables in Table 1 are reasonable, indicating that the variables are approximately normally distributed. An inspection of the raw means for language achievement across time reveals that there is no progress in the first two grades but that progress is made in language during the final years of secondary school. An inspection of the means of the school engagement variables shows a general decline in all the attitudinal variables across secondary school. In a next step we will model growth curves to the data in order to gain more insight into the nature (or form) of the developmental processes of language and school engagement.

### *Overall Development of School Engagement and Language Achievement: Univariate Latent Growth Models (Hypothesis 1)*

To gain more insight into the development of language achievement and the development of school engagement, we fitted linear and quadratic growth curves to the data. We call these univariate growth curve models because we investigated the development of one variable at a time. All growth curve models assumed that the measurement errors were heteroskedastic and uncorrelated across time.

For each outcome we fitted linear (Model 1 in Table 2) and quadratic (Model 2 in Table 2) grow curves separately. Table 2 shows the series of nested univariate latent growth curve models.

[3] At some occasions the kurtosis reached high values (see Table 1, values above 1), meaning that the distribution is peaked at some occasions. As kurtosis is a less important contributor than skewness to nonnormality (Cole & Green, 1992), it can be assumed that the variables are approximately normally distributed.

**Table 2.** Nested Model Comparisons and Model Fit Indices in Univariate Latent Growth Curve Analyses of Language Achievement, Interest in Learning Tasks, Relationship with Teachers, Attitude toward Homework, and Effort

| | $\chi^2$ | *df* | Model Comparison | $\Delta\chi^2$ | $\Delta df$ | CFI | TLI | RMSEA |
|---|---|---|---|---|---|---|---|---|
| *Language achievement* | | | | | | | | |
| Model 1: Linear growth | 152.48** | 10 | | | | 0.97 | 0.97 | 0.08 |
| Model 2: Quadratic growth | 10.32 | 6 | Model 1 vs. Model 2 | 142.53** | 4 | 0.99 | 0.99 | 0.02 |
| *Interest in learning tasks* | | | | | | | | |
| Model 1: Linear growth | 276.58** | 5 | | | | 0.84 | 0.81 | 0.16 |
| Model 2: Quadratic growth | 8.08* | 1 | Model 1 vs. Model 2 | 268.45** | 4 | 0.99 | 0.98 | 0.06 |
| *Relationship with teachers* | | | | | | | | |
| Model 1: Linear growth | 40.65** | 5 | | | | 0.98 | 0.98 | 0.06 |
| Model 2: Quadratic growth | 3.80 | 1 | Model 1 vs. Model 2 | 36.85** | 4 | 0.99 | 0.99 | 0.04 |
| *Attitude toward homework* | | | | | | | | |
| Model 1: Linear growth | 53.69** | 5 | | | | 0.97 | 0.97 | 0.07 |
| Model 2: Quadratic growth | 36.94** | 1 | Model 1 vs. Model 2 | 16.75* | 4 | 0.98 | 0.88 | 0.13 |
| *Effort* | | | | | | | | |
| Model 1: Linear growth | 8.14 | 5 | | | | 0.98 | 0.98 | 0.02 |
| Model 2: Quadratic growth | 1.84 | 1 | Model 1 vs. Model 2 | 6.30 | 4 | 0.99 | 0.98 | 0.02 |

*$p < .01$, **$p < .001$.

As these two models are nested, we can use the $\chi^2$ difference tests to compare both models and to decide which model (i.e., linear or quadratic growth curve model) fits the data better. The results of the $\chi^2$ difference tests (see Table 2; $\Delta\chi^2$) showed that a quadratic growth curve clearly fitted the data better than a linear growth curve did for each of the dependent variables except for the variable "effort" ($\Delta\chi^2$ is not significant). But because the $\chi^2$ difference tests, like $\chi^2$ tests, are sensitive to sample size, even a small difference may be significant in a large sample, such as is the case in our study (Barrett, 2007; Kaplan, 2000). Therefore, we also took into account other fit indices. The CFI, TLI, and RMSEA all indicate a better fit of the quadratic growth curves for language achievement, for the interest in learning tasks, and for the relationship with teachers (see Table 2). In addition, the overall $\chi^2$ test of model fit (see Table 2, $\chi^2$) was not significant—notwithstanding the large sample size—indicating a good model fit of the quadratic growth curves above the linear growth curves of language achievement and relationship with teachers. Based on the overall $\chi^2$ test, the $\Delta\chi^2$ tests, and the different fit indices, we conclude that quadratic growth curves represent the rate of change better than linear growth curves for language achievement, interest in learning tasks, and relationship with teachers. However, for effort and for the attitudes toward homework, we retained the linear growth curve models. Even though the $\chi^2$ difference test indicated a better fit of the quadratic model for the attitudes toward homework, the CFI, TLI, and RMSEA clearly indicated a superior fit of the linear model. For effort, the fit indices indicated that both linear and quadratic models were equivalent. Therefore, we decided to retain the simpler linear model.

Table 3 gives the means, variances, and covariances between the growth factors for language achievement, interest in learning tasks, relationship with teachers, attitude toward homework, and effort.

For language achievement, we found an initial decline (i.e., negative slope –.07; see Table 3), but this decline did not persist (i.e., positive quadratic term .02; see Table 3). The positive quadratic term indicates that students, after an initial decline, make progress in language achievement in secondary school. Furthermore, as expected the results showed a decline in the four school engagement variables (i.e., negative slopes; see Table 3). For interest in learning tasks and relationship with teachers this decline did not persist (i.e., positive quadratic term .03 and .01, respectively; see Table 3), resulting in a deceleration of the decline with time. For effort and attitudes toward homework, in contrast, the decline did not become smaller but instead persists with time.

**Table 3.** Univariate Latent Growth Curve Models and the Effect of Gender

| | Language | | Interest in Learning Tasks | | Relationship with Teachers | | Attitude toward Homework | | Effort | |
|---|---|---|---|---|---|---|---|---|---|---|
| | β | *SE* | β | *SE* | β | *SE* | β | *SE* | β | *SE* |
| Means | | | | | | | | | | |
| I | –0.07 | 0.05 | 3.99**** | 0.04 | 3.93**** | 0.04 | 4.39**** | 0.02 | 3.24**** | 0.04 |
| S | –0.07**** | 0.02 | –0.35**** | 0.02 | –0.12**** | 0.02 | –0.10**** | 0.00 | –0.06**** | 0.00 |
| Q | 0.02**** | 0.00 | 0.03**** | 0.00 | 0.01** | 0.00 | | | | |
| Variances | | | | | | | | | | |
| I | 0.30**** | 0.02 | 0.65**** | 0.05 | 0.51**** | 0.06 | 0.30**** | 0.02 | 0.13**** | 0.02 |
| S | 0.03**** | 0.01 | 0.18**** | 0.02 | 0.12**** | 0.02 | 0.01**** | 0.00 | 0.01**** | 0.00 |
| Q | 0.00*** | 0.00 | 0.00**** | 0.00 | 0.00**** | 0.00 | | | | |
| Covariances | | | | | | | | | | |
| I-S | –0.02 | 0.01 | –0.26**** | 0.03 | –0.19**** | 0.03 | –0.03**** | 0.00 | –0.01 | 0.00 |
| I-Q | 0.00 | 0.00 | 0.03**** | 0.00 | 0.02**** | 0.00 | | | | |
| S-Q | –0.00**** | 0.001 | –0.02**** | 0.00 | –0.01**** | 0.00 | | | | |
| Gender[a] on | | | | | | | | | | |
| I | –0.05 | 0.10 | 0.10 | 0.07 | 0.13* | 0.07 | 0.18**** | 0.04 | 0.03 | 0.06 |
| S | 0.17**** | 0.02 | –0.10** | 0.04 | –0.04 | 0.04 | 0.02*** | 0.01 | 0.04*** | 0.01 |
| Q | –0.02**** | 0.00 | 0.02*** | 0.00 | 0.01 | 0.01 | | | | |

*$p < .10$, **$p < .05$, ***$p < .01$, ****$p < .001$.

I = intercept, S = slope, Q = quadratic term.

[a]Gender is dummy coded with 0 for boys and 1 for girls.

### *Gender Differences in the Development of School Engagement and Language Achievement (Hypothesis 2)*

We expected that boys would show less progress in language achievement and a steeper decline in school engagement. Table 3 shows the effects of gender on the growth factors (i.e., intercept, slope, and quadratic term) of language achievement and on the growth factors of school engagement. As shown in Figure 2, girls and boys started secondary school with comparable language scores (i.e., non-significant effect of gender on intercept –.05; $d$ = –0.09),[4] but girls made initially more progress than boys did (i.e., positive effect of gender on slope .17; $d$ = 1.15). Boys, however, caught up with girls by the end of secondary school thanks to their acceleration in growth rate (i.e., negative effect of gender on the quadratic term –.02; $d$ = –0.76). According to Cohen's (1988) criteria for effect sizes, these gender differences in instantaneous rate of change and curvature in language achievement are large.

The gender differences on the interest in learning tasks and the relationship with teachers showed similar patterns (Figures 3 and 4). Girls and boys started secondary school with similar interest in learning tasks (see nonsignificant effect of gender on intercept .10; $d$ = 0.13) and with a similar positive relationship with teachers (see marginally significant effect of gender on intercept .13; $d$ = 0.18). However, girls' interest in learning tasks initially decreased to a larger extent than that of boys (negative effect of gender on slope –.10; $d$ = –0.24), but girls recovered more than boys (i.e., positive effect of gender on quadratic term .02; $d$ = 0.27) by the end of secondary school. These gender differences are small. We found no significant gender differences in the instantaneous rate of change and in the curvature of the relationship with teachers.

With regard to the attitude toward learning task and the effort for language, we found that girls started secondary school with a significantly higher attitude toward learning tasks (positive effect of gender on intercept .18; $d$ = 0.32), while girls and boys did not significantly differ in the effort for language according to the language teacher (nonsignificant effect of gender on intercept .03; $d$ = 0.09) at the start of secondary school (Figures 5 and 6). As expected, boys' attitudes toward homework (see positive effect of gender on slope .02; $d$ = 0.16) and boys' effort (see positive effect of gender on slope .04; $d$ = 0.68) declined to a larger extent than that of girls, resulting in a widening of the gender gap across secondary school in both measures. The

[4] The effect sizes are computed by dividing the gender coefficient in Table 3, which represents the gender difference on the intercept, slope, and quadratic term of an outcome variable, by the standard deviation of, respectively, the intercept, slope, and quadratic term of the outcome variable.

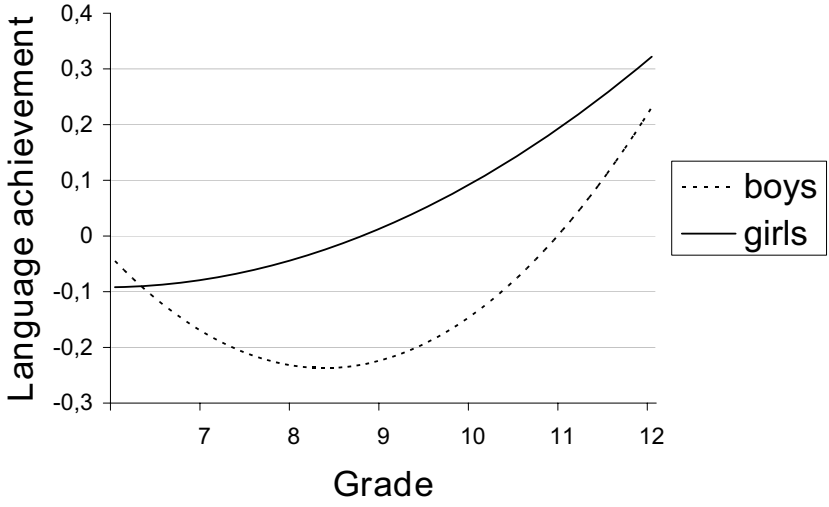



**Figure 2.** Latent Growth Curve Model for Boys' and Girls' Language Achievement across Grades 7–12.

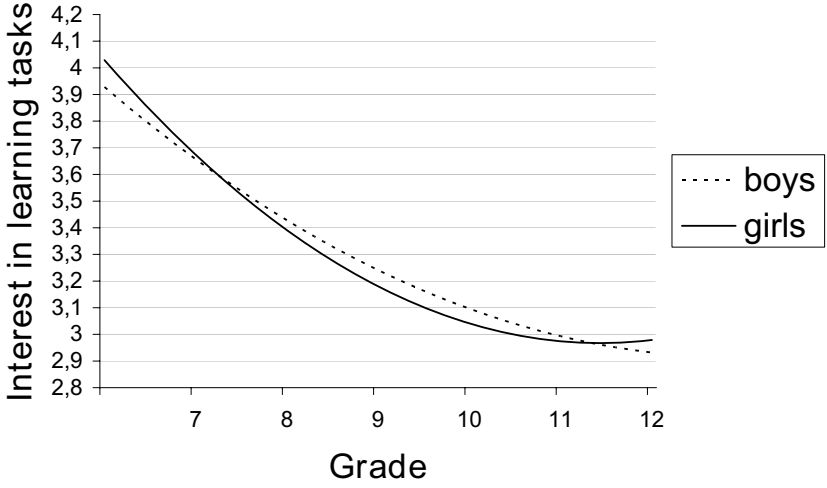



**Figure 3.** Latent Growth Curve Model for Boys' and Girls' Interest in Learning Tasks across Grades 7–12.

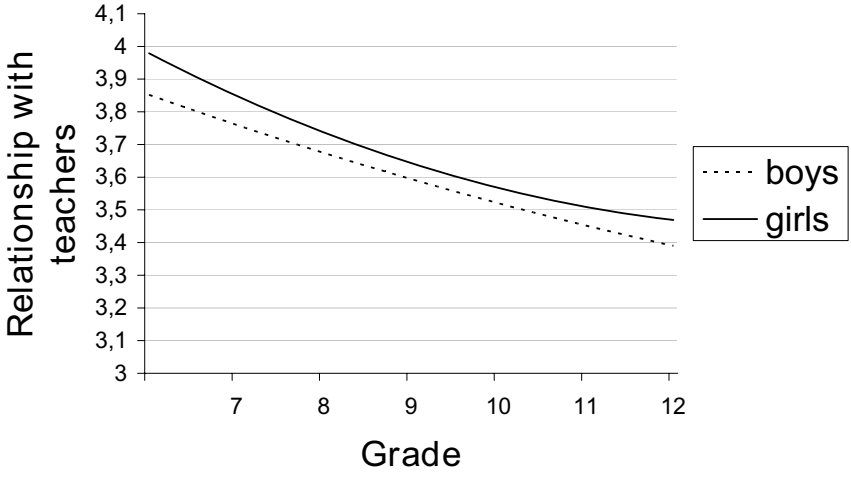



**Figure 4.** Latent Growth Curve Model for Boys' and Girls' Relationship with Teachers across Grades 7–12.

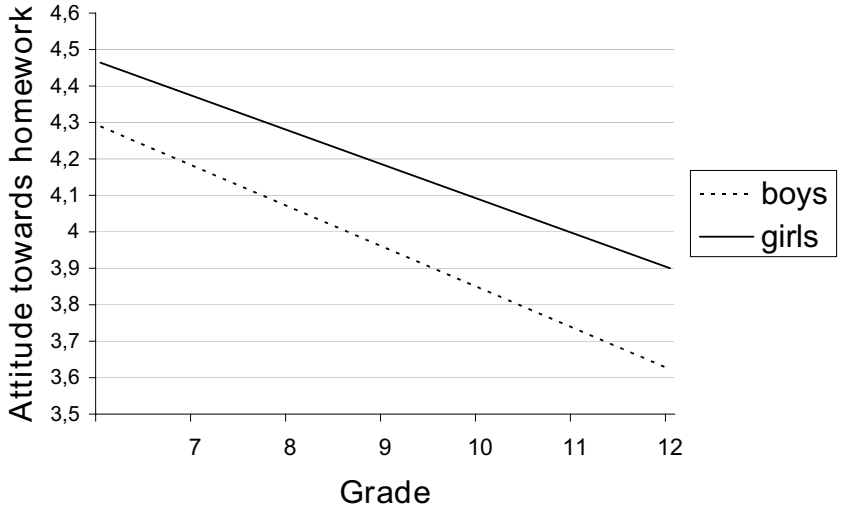



**Figure 5.** Latent Growth Curve Model for Boys' and Girls' Attitude toward Homework across Grades 7–12.

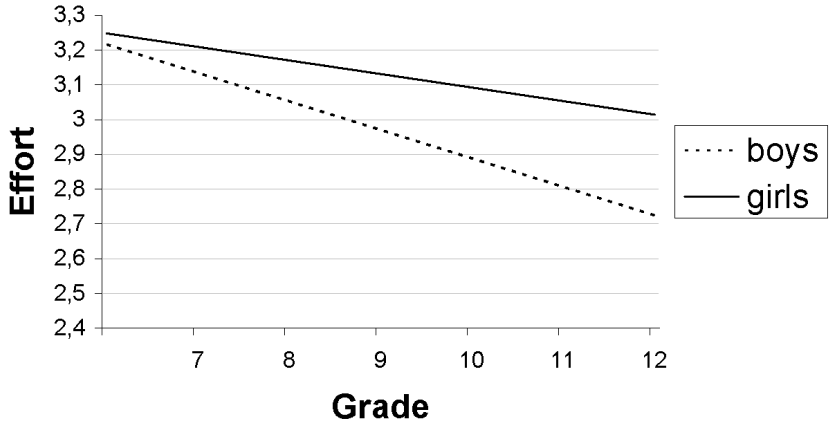



**Figure 6.** Latent Growth Curve Model for Boys' and Girls' Effort across Grades 7–12.

effect sizes of the gender differences in school engagement measures appear to be, on the whole, smaller than those in language achievement.

### *Overall Relationship between the Development of School Engagement and Language Achievement: Multivariate Latent Growth Curve Models (Hypothesis 3)*

In a next step, we used multivariate growth curve models that enable us to estimate the development of language and school engagement simultaneously (see Figure 1). The results are shown in Table 4.

First, we examined the relationships between the starting levels of each of the four school engagement variables and the starting level of language achievement. Only for the effort for language were the starting levels significantly related (see Table 4, positive effect of intercept on intercept, .50, $p < .001$). This means that students who put a lot of effort into language also score higher in language at the beginning of secondary school. For the interest in learning tasks, the relationship with teachers, and the attitude toward learning task, no such relationship was found with language achievement (respectively, see nonsignificant parameter estimates –.03, –.01, and –.05 in Table 4).

Second, we investigated the relationship between the rate of change of each of the four school engagement variables and the rate of change of language achievement. The results in Table 4 show that the development

**Table 4.** Associations between the Growth Factors of Interest in Learning Task, Relationship with Teachers, Attitude toward Homework, and Effort and the Growth Factors of Language Achievement

| | Language Achievement | | | | | |
|---|---|---|---|---|---|---|
| | I | | S | | Q | |
| | β | *SE* | β | *SE* | β | *SE* |
| *Interest in learning tasks* | | | | | | |
| I | –0.03 | 0.02 | 0.04** | 0.02 | –0.01* | 0.00 |
| S | | | 0.08* | 0.03 | | |
| Q | | | | | 0.09* | 0.04 |
| *Relationship with teachers* | | | | | | |
| I | –0.01 | 0.03 | 0.09*** | 0.02 | –0.01*** | 0.00 |
| S | | | 0.20*** | 0.05 | | |
| Q | | | | | 0.21** | 0.07 |
| *Attitude toward homework* | | | | | | |
| I | –0.05 | 0.04 | 0.07*** | 0.02 | –0.01** | 0.00 |
| S | | | 0.07** | 0.03 | | |
| *Effort* | | | | | | |
| I | 0.50*** | 0.11 | 0.04 | 0.03 | 0.00 | 0.01 |
| S | | | 0.28*** | 0.08 | | |

*$p < .05$, **$p < .01$, ***$p < .001$.

I = intercept, S = slope, Q = quadratic term.

trajectories of language achievement and of all the four school engagement variables are significantly positively related (interest in learning tasks: see positive effect of linear term on linear term .04, $p < .01$ and quadratric term on quadratic term .09, $p < .05$; relationship with teachers: see positive effect of linear term on linear term .20, $p < .001$ and quadratric term on quadratic term .21, $p < .01$; attitude toward homework: see positive effect of linear term on linear term .07, $p < .01$; effort for language: see positive effect of linear term on linear term .28, $p < .001$). This result indicates that students who showed less of an instantaneous decline in school engagement have higher instantaneous learning rates in language and that students whose decline in school engagement leveled of

by the end of secondary education showed an acceleration in their learning rates in language by the end of secondary education.

### *The Relationship between the Development of School Engagement and of Language Achievement for Boys and Girls (Hypothesis 4)*

Finally, we examined whether the associations between the growth parameters of school engagement and of language achievement differed between boys and girls. To address this question, we will compare two nested models. In a first model we constrained the associations between the growth curve parameters of school engagement and the growth curve parameters of language achievement to be equal for boys and girls. In a second model these associations are set to be different for boys and girls. By comparing the fit indices and the $\chi^2$ difference test of these two models, we can determine which model fits best. We hypothesized that the associations are stronger for boys than for girls (Hypothesis 4).

First, we examined whether the relationships between the starting levels of each of the four school engagement variables and the starting level of language achievement were different for boys than for girls. This was not the case.

Second, we examined whether the relationship between the rate of change of each of the four school engagement variables and the rate of change of language achievement was different for boys than for girls. Again, no evidence was found for gender differences.

We did, however, find significant gender differences in the association between the starting levels of school engagement (i.e., the interest in learning tasks and the relationship with teachers) and changes in language achievement. The association between the starting level of the relationship with teachers and the rate of change (linear as well as quadratic term) in language achievement was significant only for boys (linear term $\Delta\chi^2$ [1, $N$ = 2270] = 6.95, $p < .05$: for boys, $\beta = 0.12$, $SE = 0.03$; for girls, $\beta = 0.03$, $SE$ = 0.02; quadratic term $\Delta\chi^2$ [1, $N$ = 2270] = 5.87, $p < .05$: for boys, $\beta = -0.01$, $SE = 0.004$; for girls, $\beta = -0.003$, $SE = 0.003$). Similarly, only for boys was the starting level of the interest in learning tasks positively associated with the instantaneous rate of change in language achievement, $\Delta\chi^2$ (1, $N$ = 2270) = 4.10, $p < .05$ (linear term: for boys $\beta = -0.01$, $SE = 0.003$, for girls $\beta$ = –0.002, $SE = 0.002$). This means that boys who started secondary school with a positive relationship with teachers and a higher interest in learning tasks made initially more progress but showed less acceleration in the growth rate in language achievement than boys who started secondary school with a less positive relationship with teachers and with lower levels

of the interest in learning tasks. Similarly, boys who started secondary school with a higher interest in learning tasks made more progress initially in language achievement than boys who started secondary school with lower levels of the interest in learning tasks.

## Discussion

### *Summary of the Results*

Overall, we found support for our hypotheses. First, as expected, we found an overall decline in school engagement across secondary school and an overall increase in language, although there was stagnation during the first two years of secondary school. We will return to this result later.

Second, we found a stronger decline for boys in attitude toward homework and effort for language. Although boys showed less positive relationships with teachers across secondary school than girls, no gender differences in the decline of the relationship with teachers were found. In addition, boys' decline in their interest in learning tasks decelerated less than that of girls. With regard to language achievement, we found that boys showed less learning gains than girls but only during the first two years of secondary education.

Third, we found support for the hypothesis that the development of language achievement was positively associated with the development of school engagement (Hypothesis 3). The most important finding was the co-occurrence of the decline in school engagement and the increase in learning gains in language. In other words, given that all students showed a decline in school engagement and an increase in language abilities across secondary school, we found that students who showed a smaller decline in school engagement also showed a higher increase in language.

Finally, we found support for Hypothesis 4. The associations between the development of language and emotional (cf. Fredricks et al., 2004) school engagement variables were significant only for boys, indicating that school engagement is a stronger predictor of boys' than of girls' achievement.

### *Gender Differences in the Development of School Engagement and Language Achievement: Possible Explanations and Underlying Mechanisms*

A first possible explanation of why boys showed less positive developmental trajectories in school (behavioral) engagement and language achievement may be that they experience more problems with the transition between ele-

mentary and secondary school. The beginning of secondary school indeed marks the transition period between elementary and secondary school. It is during this period that we observed stagnation in their learning of language and a decline in school engagement. According to the stage-environment fit hypothesis of Eccles and Midgley (1989), there is a misfit between the characteristics of the secondary school system and the developmental needs of adolescents, and this misfit may be stronger for adolescent boys than for adolescent girls. For instance, secondary schools can be characterized by a higher focus on discipline and teacher control compared to elementary school, and this difference may be more detrimental to the engagement and the achievement of boys. Boys' higher prevalence of disruptive behavior in the classroom may be a reflection of boys' lower compliance to school rules and their relative lack of acceptance of teacher control and discipline (Johnson, McGue & Iacono, 2005; Stowe et al., 2000). Some authors have indeed argued that secondary schools and classrooms are feminized environments, meaning that teachers' attitudes, classroom rules, and learning tasks favor girls (Newkirk, 2002). However, these are only suggestions. Based on our research, we cannot make any causal assumptions between school transition and the lower engagement and language achievement of boys. More research is needed to uncover exactly which characteristics of the educational environment show a larger mismatch with the educational needs of boys than with those of girls.

A second possible explanation may be differences in maturation rates between boys and girls. The finding that boys' growth in language achievement was less than that of girls in the beginning of secondary school but that they recovered and caught up with girls by the end of secondary school seems to support such a maturation hypothesis. Although we cannot rule out the possibility that gender differences in maturation rates plays a role in the development of gender differences in achievement and engagement, we would like to encourage more researchers to disentangle the confounding effect of transition and age on the development of engagement and achievement (Simmons and Blyth, 1987). As Eccles et al. (1993, p. 90) stated, "Few developmental periods [adolescence] are characterized by so many changes at so many different levels—changes due to pubertal development, social role definition, cognitive development, school transitions, and the emergence of sexuality." Thus, multiple causal mechanisms may be at work.

Another important result that needs some further explanation is that we only found a larger decline for boys in the behavioral engagement variables but not in the emotional engagement variables. According to Fredricks et al. (2004), it is not unlikely that these two types of school engagement show different developmental patterns (and gender differences). The results of the present study indeed showed that gender differences are more important in

behavioral engagement than in emotional engagement. However, we modeled the developmental trajectories of emotional and behavioral engagement in a different way. We modeled the trajectories of emotional engagement by a quadratic growth curve, whereas we modeled those of behavioral engagement by a linear curve. These two different ways of modeling the developmental trajectories may influence, mask, or expose how gender differences become apparent. The results showed that by modeling more complex growth curves (i.e., emotional engagement), more complex gender differences emerged. Therefore, more research is needed to determine whether gender differences in behavioral engagement are more important.

### *Recommendations for Future Research*

An interesting future line of research would be to investigate the effects of changes in school characteristics on changes in school engagement directly. This means that longitudinal data need to be collected not only at the student level but also at the school level in order to investigate the associations between changes in student engagement and changes in school characteristics. Although previous research has focused on the effects of school characteristics on school engagement (e.g., Finn & Voelkl, 1993), little research has addressed the relationship between school characteristics and student engagement from a longitudinal perspective.

The results of the present study support Finn's (1989, 1993) theory of school dropout that states that students who gradually become less engaged with schooling show a decline in their (language) achievement. Repeatedly being confronted with low achievement may lead to feelings of incompetence accompanied by disengagement from school, and this may eventually lead to school dropout. It would be interesting to look at associations between the decline in school engagement and other factors besides achievement, such as retention rates and school dropout. In addition, these examinations should not be limited to gender differences. It is also important to investigate differences between students with different ethnic and socioeconomic backgrounds, as these groups may show different developmental trajectories that may be linked to differences in school dropout rates.

Finally, our results are limited to the period of secondary school. At the beginning of secondary school, however, boys showed less positive attitudes toward homework. This indicates that gender differences in school engagement emerge prior to secondary school. It is thus important to also examine what is happening in primary education to understand why certain boys end primary education less engaged than girls. Future research should concentrate on gender differences in the development of engagement in pri-

mary education and on what factors contribute to the decline of engagement, especially for boys.

### *Limitations*

Despite several strengths of the present study such as the large data set, the longitudinal perspective on the association between school engagement and achievement, and the special focus on gender differences, our study also has some limitations. One of the limitations concerns the fact that the measures of school engagement, except effort, are not subject specific. Interest in learning tasks, relationship with teachers, and attitude toward homework are general engagement variables and do not specifically refer to the subject of language. If we had measured subject specific school engagement variables such as the interest in language instead of the interest in learning tasks, then a stronger relationship with language achievement might have been found.

A second possible limitation concerns our measure of student effort. We used only one item to measure effort. In addition, the teachers' reported effort for language might be confounded with teachers' expectancies and beliefs about the cognitive abilities of the students. Although we obtained meaningful results with this variable that coincide with the results of studies that used more reliable instruments to measure effort for language (Watt, 2004b), it should be kept in mind that this measure may suffer from lower reliability and validity. The fact that we only found a significant relationship between the starting levels of the effort for language and language achievement may be confounded by the expectation of teachers. Therefore, we recommend using self-reported effort in future research, although this measure may also not be entirely free of measurement error (e.g., students may overestimate or underestimate their effort for language).

Finally, the longitudinal data in the present study is restricted to students from the academic and technical track that were not retained during secondary school. This may result in a more selective sample for boys because (a) more girls than boys follow the academic track and (b) more boys than girls are retained in secondary school. This may have attenuated the size of the gender differences in language achievement and in school engagement, but we have no theoretical reasons to assume that this would affect the associations between language achievement and school engagement.

### *Conclusions*

Our main research question was whether boys become more disengaged with schooling than girls over time and whether this negative development

is accompanied by a decline in their learning rates in language. The present study provided evidence for this hypothesis. Boys in particular get disengaged from school as they put less effort into their work and develop less positive attitudes toward homework compared to girls, and this negative development seemed to be reflected in boys' learning rates in language. As such, our study identifies one of the crucial factors that may lie at the base of the persisting gender gap in language achievement. Our findings are important in that they demonstrate how the development of school engagement and achievement are related over time and how this shows different patterns for boys and girls. Our results provide suggestions for practitioners and intervention programs that are aimed at pinpointing and ameliorating the gender gap in school performance of poor-achieving youths.

## References

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of Educational Research, 64,* 287–309.

Arnot, M., David, M., & Weiner, G. (1999). *Closing the gender gap: Postwar education and social change.* Cambridge, UK: Polity.

Barnett, R. C., & Rivers, C. (2006). The boy crisis—fact or myth? *Teachers College Record,* October 2.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Structural Equation Modeling, 42*(5), 815–824.

Burgess, S., McConnell, B., Propper, C., & Wilson, D. (2004). Girls rock, boys roll: An analysis of the age 14–16 gender gap in English schools. *Scottish Journal of Political Economy, 51,* 209–229.

Clark, A., & Trafford, J. (1995). Boys into modern languages: An investigation of the discrepancy in attitudes and performance between boys and girls in modern languages. *Gender & Education, 7,* 315–325.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences.* Hillsdale, NJ: Erlbaum.

Cole, N. S. (1997). *The ETS gender study: How females and males perform in educational settings.* Princeton, NJ: Educational Testing Service.

Cole, T. J., & Green, P. J. (1992). Smoothing reference centiles curves: The LMS method and penalized likelihood. *Statistics in Medicine, 11,* 1305–1319.

Coley, R. J. (2003). *Growth in school revisited: Achievement gains from the fourth to the eighth grade.* Princeton, NJ: Educational Testing Service.

Davies, J., & Brember, I. (2001). The closing gender gap in attitudes between boys and girls: A 5 year longitudinal study. *Educational Psychology, 21,* 103–114.

Downey, D. B., & Yuan, A. S. V. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *Sociological Quarterly, 46,* 299–321.

Eccles, J. S., & Midgley, C. M. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (pp. 139–186). San Diego: Academic Press.

Eccles, J. S., Midgley, C., Wigfield, M., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in school and in families. *American Psychologist, 48,* 90–101.

Engels, N., Aelterman, A., Van Petegem, K., & Schepens, A. (2004). Factors which influence the well-being of pupils in Flemish secondary schools. *Educational Studies, 30,* 127–143.

Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling, 6,* 56–83.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59,* 117–142.

Finn, J. D. (1993). *School engagement and students at risk.* Washington, DC: National Center for Education Statistics.

Finn, J. D., & Voelkl, K. E. (1993). School characteristics related to student engagement. *Journal of Negro Education, 62,* 249–268.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of evidence. *Review of Educational Research, 74,* 59–109.

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence. *Developmental Psychology, 38*(4), 519–533.

Freudenthaler, H., Spinath, B., & Neubauer, A. (2008). Predicting school achievement in boys and girls. *European Journal of Personality, 22,* 231–245.

Hill, P. W., & Russell, V. J. (1999). Systemic, whole-school reform of the middle years of schooling. In R. J. Bosker, B. P. M. Creemers, & S. Stringfield (Eds.), *Enhancing educational excellence, equity and efficiency: Evidence from evaluations of systems and schools in change* (pp. 167–196). Dordrecht: Kluwer Academic.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Hyde, J. S., & Kling, K. C. (2001). Women, motivation, and achievement. *Psychology of Woman Quarterly, 25,* 364–378.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73,* 509–527.

Janssen, P. (1982). *Vragenlijst studiebeleving* [Study experience questionnaire]. Leuven: K.U.Leuven, Afdeling Psychodiagnostiek en Psychologische Begeleiding, Centrum voor Schoolpsychologie.

Johnson, W., McGue, M., & Iacono, W. G. (2005). Disruptive behavior and school grades: Genetic and environmental relations in 11-years-olds. *Journal of Educational Psychology, 97,* 391–405.

Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions.* Thousand Oaks, CA: Sage.

Kiplinger, V. L. (2004). *Longitudinal study of student growth in reading, writing, mathematics achievement in thirty-six Colorado public school district, Phase II: Final Report.* Retrieved on June 5, 2009, from http://matrix10.asd20.org/lssg/phase2/Phase2Report.pdf.

Kolaweski-Jones, L., & Duncan, G. J. (1999). The structure of achievement and behaviour across middle childhood. *Child Development, 70,* 930–943.

Lamb, S. (1997). Gender differences in mathematics participation: An Australian perspective. *Educational Studies, 23,* 105–126.

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education, 34*(1), 89–109.

Maruyama, G. M. (1998). *Basics of structural equation modeling.* Thousand Oaks, CA: Sage.

Mead, S. (2006). *The truth about boys and girls.* Washington, DC: Education Sector.

Muthén , B. O., & Khoo, S.-T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, 10,* 73–101.

Muthén, L. K., & Muthén, B. O. (1998–2004). *Mplus user's guide.* Los Angeles, CA: Muthén and Muthén.

Newkirk, T. (2002). *Misreading masculinity: Boys, literacy, and popular culture.* Portsmouth, NH: Heinemann.

Olson, L. (2002). Detachment starts in middle schools, study finds. *Education Week, 21,* 9.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74,* 525–556.

Sayer, A. G., & Willett, J. B. (1998). A cross-domain model for growth in adolescent alcohol expectancies. *Multivariate Behavioural Research, 33,* 509–543.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context.* Hawthorn, NY: Aldine de Gruyler.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford: Oxford University Press.

Smits, J., & Vorst, H. (1982). *Schoolvragenlijst voortgezet onderwijs* [School questionnaire secondary education]. Nijmegen, Netherlands: Berkhout Nijmegen.

Stoel, R. D., Peetsma, T. T. D., & Roeleveld, J. (2003). Relations between the development of school investment, self-confidence, and language achievement in elementary education: A multivariate latent growth curve approach. *Learning and Individual Differences, 13,* 313–333.

Stowe, R. M., Arnold, D. H., & Ortiz, C. (2000). Gender differences in the relationship of language development to disruptive behaviour and peer relationships in preschoolers. *Journal of Applied Developmental Psychology, 20,* 521–536.

Sum, A., Kirsch, I., & Taggert, R. (2002). *The twin challenges of mediocrity and inequality: Literacy in the U.S. from an international perspective.* Princeton, NJ: Policy Information Report Educational Testing Service. Available at http://www.ets.org/Media/Research/pdf/PICTWIN.pdf.

Van Damme, J., & Onghena, P. (Eds.) (2002). Educational effectiveness in secondary schools in Flanders [Special issue]. *School Effectiveness and School Improvement,* 13(4).

Van de gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2007). Impact of attitudes of peers on language achievement: Gender differences. *Journal of Educational Research, 101*(2), 78–90.

Van Schooten, E., de Glopper, K., & Stoel, R. D. (2004). Development of attitude toward reading adolescent literature and literacy reading behaviour. *Poetics, 32,* 343–386.

Walsh, M., Hickey, C., & Duffy, J. (1999). Influence of item content and stereotype situation on gender differences in mathematical problem solving. *Sex Roles, 41,* 219–240.

Watt, H. M. G. (2004a). *Boys and English: Do boys "disengage" from English through secondary school?* Paper presented at the 3rd International Biennial SELF Research Conference, Berlin, July 2004.

Watt, H. M. G. (2004b). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Development, 75,* 1556–1574.

Whitelaw, S., Milosevic, L., & Daniels, S. (2000). Gender, behaviour and achievement: A preliminary study of pupil perceptions and attitudes. *Gender & Education, 12,* 87–113.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25,* 68–81.

Wigfield, A., Harold, R. D., Freedman-Doan, C., Eccles, J. S., Suk Yoon, K., Arbetron, A. J. A., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values cross the elementary school years: A 3-year study. *Journal of Educational Psychology, 89,* 451–469.

Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (1994). *BIMAINTM2: Multiple-group IRT analysis and test maintenance for binary outcomes.* Chicago: Scientific Software International.

## Appendix 1
## Item Composition of the School-Related Attitudinal Scales

*Interest in Learning Tasks (8 Items)*

- I enjoy most of the subjects taught in this school.
- To me, many things we have to learn in school are unimportant. (–)*
- I think that I learn useful things in school.
- I am really interested in most of the subjects.
- I think it's great that I learned all sorts of things this year.
- I think that most of the subjects we are taught are very worthwhile.
- I think that I have to learn things in school that I won't ever need in future. (–)
- Personally, I find the subject matter usually interesting.

*Relationship with Teachers (10 Items)*

- I think that most of the teachers are very helpful when I have problems with school work.
- Some teachers are kinder to others than to me. (–)
- I feel at ease with most of the teachers.
- There are few teachers who help me well with my school work. (–)
- There are enough teachers who listen patiently when I ask something.
- I get on well with most of the teachers.
- There are few teachers who understand me. (–)
- Some teachers don't have the patience to explain things to me. (–)
- The teachers dislike me. (–)
- Most of the teachers treat me in a nice way.

*Attitude toward Homework (5 Items)*

- When I have homework, I put it off for as long as possible before I start. (–)
- When I have homework, I start as soon as possible.
- I usually start doing my homework of my own accord.
- When I want to do something nice, I still complete my homework first.
- At home, I only start doing my homework when I am told to do so. (–)

* Items that have a negative contribution to the scale are indicated by (–).