**Mathematical Thinking, Learning and Performance:**

Reed, H.C.

2014

# 6

# Keeping on track: Performance profiles of low performers in academic educational tracks

Helen C. Reed, Floryt van Wesel, Carolijn Ouwehand and Jelle Jolles

*British Educational Research Journal (2014), doi: 10.1002/berj.3132*

## ABSTRACT

In countries with high differentiation between academic and vocational education, an individual's future prospects are strongly determined by the educational track to which he or she is assigned. This large-scale, cross-sectional study focuses on low-performing students in academic tracks who face being moved to a vocational track. If more is understood about these students, measures could be taken to improve their performance and keep them within academic education. The study investigates performance patterns in academic tracks in the first three years of secondary school in the Netherlands. By identifying patterns that reveal how competence levels in different domains are related at different stages of development and by comparing low performers with other students, the study sheds light on individual and educational aspects that may be amenable to intervention. School grades were analysed for 1,596 students. School performance was found to reflect three domains - *languages* (language of schooling and modern foreign languages), *social studies* and *science and math* - that appear to interact in a process of co-construction. General language skills were robustly related to performance in other domains - particularly *social studies* - throughout the first three years of secondary school. By comparison, proficiency specifically in the language of schooling was less strongly related to *social studies* and *science and math* performance after the first year. Suggestions are given as to how educators and curriculum developers could use these insights to accommodate individual and developmental differences and to develop learning materials that may help low performers keep on track.


**Keywords:** school performance; low performers; language skills; language of schooling

## INTRODUCTION

In countries with educational systems that are stratified according to general scholastic ability, an individual's future occupational status and employment prospects are strongly determined by the educational track to which he or she is assigned at school. This is particularly so where there is high differentiation between academic and vocational education. In contrast to those in academic tracks, individuals in vocational tracks often have no direct access to higher education or professional employment and attain lower occupational status (Andersen & Van de Werfhorst, 2010; Shavit & Müller, 2000; Wolbers, 2007). Thereafter, those graduating from lower educational tracks with lower-level skills face a future with lower wages and higher levels of unemployment (OECD, 2012a; Oesch, 2010).

In stratified systems, students are often placed in designated educational tracks from early on in secondary school, with initial track allocation depending on performance in the final years of primary school and subsequent movement between tracks depending on ongoing performance in the first years of secondary school. In consequence, students who perform poorly in academic tracks may be moved to a vocational track. This considerably impacts their future prospects in an economic climate where higher educational levels are needed to compete in an increasingly international labour market (OECD, 2012a). Next to reduced access to higher education, vocational tracks can further inhibit educational attainment of these students in several ways: through depriving them of the motivational benefits of being in an academically successful milieu; through a more restricted curriculum and lower level of intellectual complexity; and through lowering their expectations and aspirations (Shavit & Müller, 2000). In highly differentiated systems, being moved out of academic tracks also often involves changing schools, which is even more unsettling for the individual concerned. Altogether, these far-reaching consequences make it important to understand as much as possible about low-performing students at risk of dropping out of academic tracks. It may then be possible to take specific measures to improve their performance and keep them within academic education.

Many factors are known to influence educational outcomes, including self-concept and beliefs (e.g., Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011; Marsh & Martin, 2011), motivation and engagement (e.g., Cleary & Chen, 2009; Steinmayr & Spinath, 2009), student-teacher relationships (e.g., Konishi, Hymel, Zumbo, & Li, 2010; Lessard, Poirier, & Fortin, 2010), health (e.g., Dewald, Meijer, Oort, Kerkhof, & Bögels, 2010; Shore et al., 2008; Vinciullo & Bradley, 2009), ethnic and economic background (e.g., Kingdon & Cassen, 2010; Wiggan, 2007), and teachers' stereotypes and expectations in relation to ethnic and economic background (e.g., McKown & Weinstein, 2008; Rubie-Davies, Hattie, & Hamilton, 2006; Wiggan, 2007). Nonetheless, with respect to decision-making about students' educational placement and progress, it is actual ongoing performance that plays a central role. To this end, exploring intrinsic patterns within school performance is a highly relevant and ecologically valid way to investigate factors underlying low performance. Specifically, performance patterns could reveal how competence levels in different domains are related at different stages of development. Insights into interactions between domains could inform curriculum development and educational methods with respect to timing, phasing, integration and delivery of learning materials and enable teachers to offer instruction that takes account of both individual and developmental differences. The present study takes a unique approach by exploring these issues across a wide range of content areas in a large-scale, cross-sectional sample covering several year cohorts within academic tracks. The insights obtained could help students who are at the cusp of the divide with vocational education to keep - literally - on track.

*School performance*

Central to the present study is the construct of school performance, for which several operationalisations exist. Multidimensional operationalisations - such as school grades - provide separate measures for individual content areas. Unidimensional operationalisations - such as grade point average (GPA) and International or European Baccalaureate total score - provide a single performance measure based on averages and/or weighted averages of multiple component scores. In addition, standardised tests can be used at key educational stages such as pre-admission to secondary or higher education. The English National Curriculum system requires such testing in core subjects at the end of primary school, while in the United States, SAT or ACT tests may be taken prior to college admission. In the Netherlands, a secondary school placement test is taken by approximately 85% of primary school pupils (http://www.government.nl/issues/education). Such tests reflect students' abilities at a fixed moment in time, however, rather than being an ongoing measure of actual performance. For the purpose of this study, the most relevant measure of school performance is school grades, which gauge ongoing performance in individual content areas (i.e., school curriculum subject areas).

A complication arises from this approach, however, because school grades are known to suffer from certain issues that can disturb the representation of students' abilities. Specifically, the quality of interaction between an individual and the situation in which a test is taken or other gradable work is produced can cause large fluctuations in actual performance (Lang & Lang, 2010; Nguyen & Ryan, 2008), while varying assessment and grading practices are known to impact the reliability of test scores as performance indicators (Allen, 2005; Bowers, 2011; McMillan, 2001). For these reasons, the present study focuses not on individual grades but on more robust underlying dimensions of school performance. To be precise, outcomes in individual content areas (i.e., grades) are reduced to a smaller number of variables on the basis of patterns of shared variance and covariance. Structural commonalities reflecting common underlying factors are extracted from the observed variables, thereby reducing the effect of measurement error and random fluctuations and more accurately estimating the underlying constructs of interest (Bollen, 2002; Kline, 2005; Loehlin, 2004). This approach is comparable to latent state-trait approaches in personality research, which distinguish between stable and consistent person characteristics (i.e., latent traits), the interaction between an individual and a particular situation, and measurement error (Steyer, Schmitt, & Eid, 1999).

The question then arises as to which underlying factors or dimensions of school performance should form the basis for investigating relationships between content areas in the present study. As a point of departure, a widely-used grouping of content areas with high face-validity for educators and researchers is assumed (see the Universal Decimal Classification scheme at http://www.udcc.org/). This recognises three commonly delimited academic domains: *languages* (i.e., literacy and literature in both the language of schooling and modern foreign languages), *social studies* (i.e., geography and history), and *science and math* (i.e., the natural sciences and mathematics). In this view, language of schooling is related specifically to the *languages* domain.

An alternative model of school performance is also feasible. Adolescent literacy research has shown that proficiency in the language of schooling can affect performance in all domains (e.g., Gomez & Gomez, 2007; Moje, 2002). Deficiency in the language of schooling is thought to underlie the academic underperformance of certain ethnic and economic groups (Snow & Biancarosa, 2003). For example, children whose native language differs from the language of schooling generally have lower performance throughout primary school and into secondary school (Ledoux, Roeleveld, Driessen, Cuppen, & Meijer, 2011). This persistent disadvantage has been found to derive from deficiencies in mastering those aspects of the language of schooling known as ‘academic language’ (e.g., Laghzaoui, 2011), which refers to the use of language for cognitively demanding communication in school-related tasks (Schleppegrell, 2004). Academic language is characterised by more elaborated, condensed and abstract ways of communication, with a rich and varied vocabulary that includes specific or technical words and a high density of content words (Laghzaoui, 2011; Schleppegrell, 2004). With regard to the present study, this suggests that performance in all domains could be structurally related to proficiency in the language of schooling. This would support an alternative model of school performance in which language of schooling plays an important role in all domains.

Furthermore, research into vocabulary acquisition and comprehension (Benelli, Belacchi, Gini, & Lucangeli, 2006; Nippold, Hegel, & Sohlberg, 1999) shows that, while knowledge of abstract (i.e., decontextualised or intangible) words continues to develop during adolescence and adulthood, there appears to be a period of salient growth during the move from late childhood to early adolescence (Nippold et al., 1999), which matches the period of the first years of secondary school. The rate at which individuals develop is highly variable, however, resulting in large differences in the ability to understand and use more abstract language. This suggests that the relationship between proficiency in the language of schooling and school performance may vary at different stages of development and may be stronger for younger adolescents during the aforementioned period of salient growth.

*Approach and hypotheses of the present study*

Within this framework, the present exploratory study investigates performance patterns within academic tracks in the first three years of secondary school within a stratified educational system. The approach taken is to examine and compare patterns of performance across academic domains and across year cohorts for low, intermediate and high performers. This could shed light on individual and educational factors differentiating low performers from better performing students that may be amenable to intervention. Similar approaches comparing students at different performance levels have been employed to investigate diverse cognitive and metacognitive factors affecting learning (e.g., Dagnino, Ballauri, Benigno, Caponetto, & Pesenti, 2013; Dermitzaki, Andreou, & Paraskeva, 2008; Rao, Moely, & Sachs, 2000). However, this is the first study to our knowledge that uses this approach to explore patterns of performance across different academic domains as well as including several year cohorts to allow investigation of developmental differences.

As a preliminary and subsidiary to addressing these main issues, it was first necessary to establish a model of school performance that could be used to investigate domain relationships. To this end, it was determined whether to model school performance in terms of the three commonly-delimited domains of *languages*, *social studies* and *science and math*, or whether to relate proficiency in the language of schooling to not only the *languages* but also other domains. On the basis of research into vocabulary development, the suitability of these models was expected to vary at different stages of development, with a stronger role for the language of schooling for younger adolescents in both the *languages* and other domains.

## METHODS

### Participants

The study took place in the Netherlands, where the educational system is characterised by high differentiation between academic and vocational education. Approximately 40% of the secondary school population is in academic tracks, which prepare more able students for higher education (i.e., college or university) and professional employment. There are two academic tracks: pre-university education (highest track) and higher general secondary education (second-highest track). Vocational tracks, by comparison, provide practical or vocational training to students for whom a higher level of education is considered unsuitable. When a student enters secondary school, he/she is placed in an academic or a vocational track according to the advice of the primary school and the results of a school placement test, if taken. Tracks may be combined during an orienting period to accommodate students for whom it is unclear which track is most suitable; at the end of the orienting period, these students continue in a single track.

The study used a clustered sample design, involving students in academic tracks from the first three years of four large secondary schools. Characteristics of these schools with regard to size, socio-economic status of the school population and percentage graduating school leavers are comparable to Dutch national averages for schools offering academic tracks (data provided by the Dutch Council for Secondary Education (www.schoolvo.nl)). For three schools, data were provided for two complete classes for each year cohort (i.e., six classes per school), and for one school, data were provided for four complete classes for each year cohort (i.e., twelve classes). One combined-track class ($N = 30$) was subsequently removed from the dataset because its grading level could not be established with certainty (see the description of Dutch grading practices in the Data section). The remaining sample ($N = 1{,}596$) comprised 411 students in Year 1 (Grade 7; age indication 12-13 years), 601 in Year 2 (Grade 8; age indication 13-14 years) and 584 in Year 3 (Grade 9; age indication 14-15 years). Regarding track composition of the sample, 663 students were in the highest academic track (i.e., pre-university education), 530 students were in the second-highest academic track (i.e., higher general secondary education) and 403 students were in a combined academic track (i.e., during the orienting period described above).

**Data**

School performance was measured using end-of-year grades, which incorporate results from multiple test-moments throughout the school year. End-of-year grades are the most relevant measure of performance for this study, as they form the basis for decisions about whether students may stay within their current educational track or are 'demoted' to a lower track. Descriptive statistics for the grades given for each content area are presented in Table 6.1. It should be noted that Dutch schools have some curricular freedom in the first years of secondary school: for example, in the present sample, not all schools teach Biology in the third year.

In Dutch secondary schools, grades range from 1 (extremely poor) to 10 (outstanding), with a pass usually set at 5.5. In addition to possible grading differences between schools, there are structural differences between educational tracks arising from differences in levels of test difficulty. Higher tracks are given more difficult tests, so that a grade obtained in a lower track is considered to correspond to a lower grade in a higher track. In the most general terms, a difference of approximately 1 point in grade value can be assumed when comparing grades across adjacent tracks. For example, a 6 in the higher general secondary track is roughly equivalent to a 5 in the pre-university track. To account for both between-tracks and between-schools differences, grades were transformed before analysis. First, corrections for structural track differences were performed according to the method described in the Appendix. Then, the corrected grades in each year cohort were group mean centred per school. Group mean centring is a transformation that removes between-groups variance of the variable(s) in question (Hox, 2002). In this case, the method reduces variance due to systematic grading differences between schools. For eleven cases with incidental missing data, missing values were imputed as the group mean (i.e., zero after group mean centring).

Table 6.1

*Descriptive statistics school grades*

| | Dutch | English | French | German | Geog | History | Biology | Physics | Chem | Math |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Year 1 (Grade 7) (*N* = 411) | | | | | |
| Mean | 6.79 | 6.76 | 6.89 | - | 6.72 | 6.84 | 6.72 | - | - | 6.42 |
| *SD* | 0.94 | 1.15 | 1.22 | - | 0.91 | 0.90 | 0.91 | - | - | 1.12 |
| Min | 3.5 | 3.5 | 3.0 | - | 3.5 | 4.0 | 3.9 | - | - | 3.0 |
| Max | 9.0 | 9.0 | 10.0 | - | 9.0 | 9.0 | 9.0 | - | - | 9.1 |
| | | | | | Year 2 (Grade 8) (*N* = 601) | | | | | |
| Mean | 6.60 | 6.79 | 6.48 | 7.14 | 6.69 | 6.81 | 6.71 | 6.50 | - | 6.37 |
| *SD* | 0.82 | 1.02 | 1.23 | 1.08 | 0.82 | 0.88 | 0.90 | 1.00 | - | 1.11 |
| Min | 4.0 | 4.0 | 3.0 | 4.0 | 4.0 | 4.0 | 4.0 | 3.0 | - | 3.0 |
| Max | 9.0 | 10.0 | 9.5 | 10.0 | 9.0 | 9.5 | 9.0 | 9.0 | - | 9.2 |
| | | | | | Year 3 (Grade 9) (*N* = 584) | | | | | |
| Mean | 6.56 | 6.72 | 6.50 | 6.66 | 6.60 | 6.54 | 6.61[a] | 6.44 | 6.76 | 6.33 |
| *SD* | 0.75 | 1.10 | 1.20 | 1.08 | 0.83 | 0.98 | 0.89 | 1.21 | 1.04 | 1.09 |
| Min | 3.5 | 3.0 | 3.0 | 3.5 | 4.0 | 4.0 | 4.0 | 3.0 | 3.5 | 3.5 |
| Max | 9.0 | 10.0 | 9.0 | 9.0 | 9.0 | 9.5 | 9.0 | 10.0 | 9.5 | 10.0 |

Note. [a] *N* = 321.

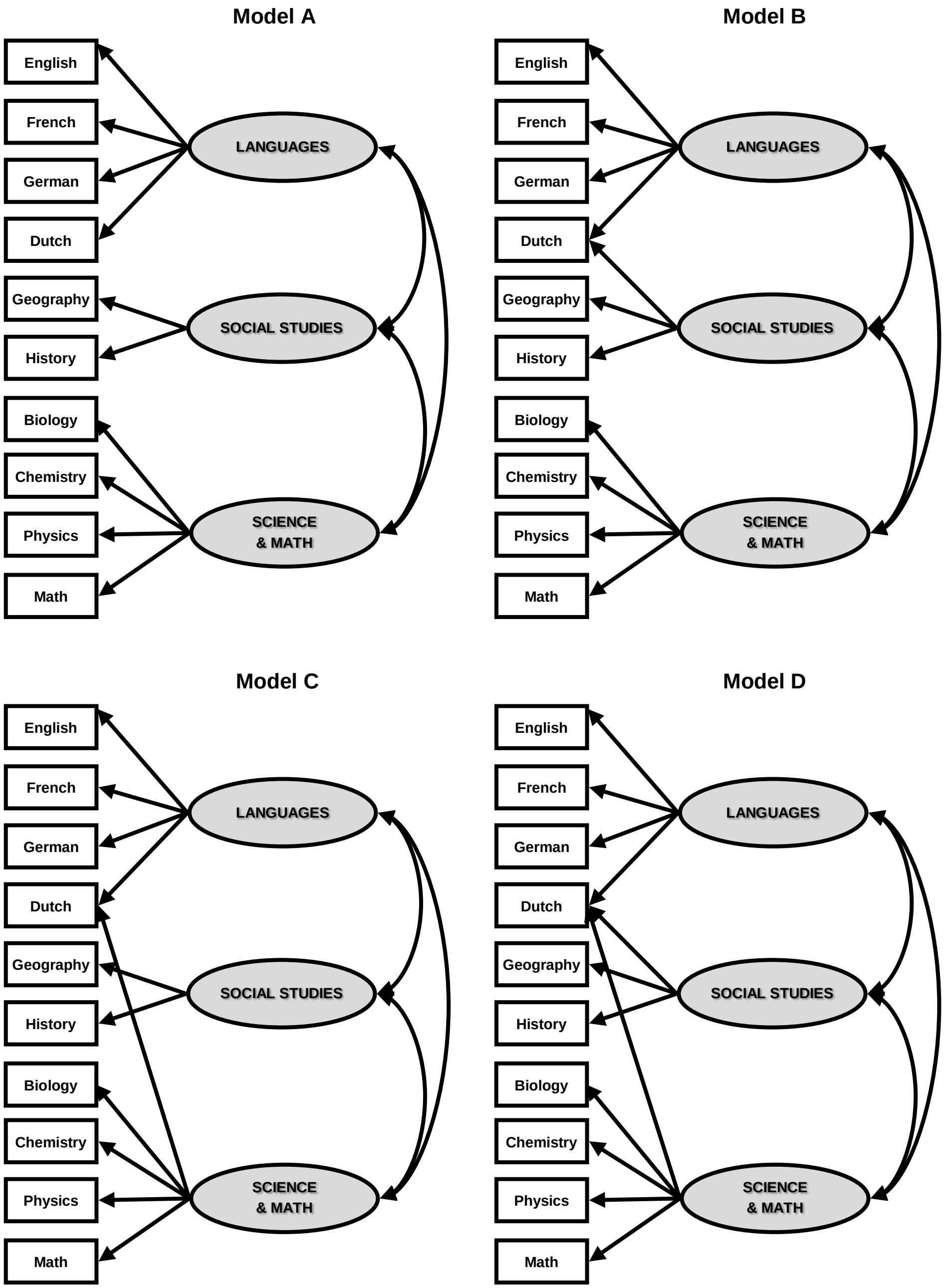
**Analysis**

*Models of school performance.* Before performance patterns could be analysed, Confirmatory Factor Analysis (CFA) established which of the hypothesised models of school performance best fit the sample data, with the transformed school grades as individual indicators. CFA uses the concept of shared variance between indicators to determine whether they tap upon some common domain (i.e., latent factor; Kline, 2005). This approach averages out a certain amount of disturbance which may arise from various sources, while explicitly accounting for measurement error in estimates of model-fit (Loehlin, 2004).

A preliminary unidimensional model tested whether a single cognitive trait underlies outcomes in all content areas. Then a model (Model A) was tested delimiting the domains *languages* (Dutch, English, French, German), *social studies* (Geography, History) and *science and math* (Biology, Chemistry, Physics, Mathematics). Subsequently, models were tested where language of schooling (Dutch) was added as an indicator for the *social studies* domain (Model B), the *science and math* domain (Model C), and for both these domains (Model D). These four models are depicted in Figure 6.1. To allow for developmental differences, models were tested for each year cohort separately. As not all content areas were taught in all school years, the models tested for each year included only those content areas taught in that year. Year 3 models were tested excluding Biology, as this was taught in only two schools in Year 3.

IBM SPSS® Amos™ 20 was used to test these models and the standard Maximum Likelihood (ML) method of estimating free parameters in structural equation models was used to asses model fit. The fit of the different models was evaluated with the Standardised Root Mean Squared Residual (SRMR), the Root Mean Squared Error of Approximation (RMSEA) and the Comparative Fit Index (CFI). With this procedure, a SRMR < .08, RMSEA < .06 and CFI > .95 together indicate good fit, while a SRMR < .10, RMSEA < .08 and CFI > .90 indicate adequate fit (Hu & Bentler, 1999). The Consistent Akaike Information Criterion (CAIC), which accounts for both parsimony and sample size, was used to select among competing models, models with lower CAIC values being preferred (Kline, 2005). In addition, model selection took account of the magnitude of the regression coefficients, with a standardised coefficient of around .20 or lower being considered too low to be of relevance (Kline, 2005).

*Low, intermediate and high performers.* Analyses of performance patterns were carried out in SPSS® Statistics 20 with an alpha level of .05. First, the CFA results were used to calculate latent factor quotient scores representing performance in each domain, following a method for calculating composite scores described by Van der Elst, Van Boxtel, Van Breukelen, and Jolles (2008). For each year cohort, the transformed grades were standardised, multiplied by their standardised regression weights (factor loadings) and aggregated for the relevant factor. The aggregated




Note. The models tested for each year include only those content areas taught in that year.

*Figure 6.1*. Hypothesised CFA models

scores were then divided by their standard deviations to form latent factor scores. Finally, for ease of interpretation, latent factor scores were linearly transformed to a scale with a mean of 100 and a *SD* of 15. The resulting quotient scores correspond conceptually to scores for each academic domain, and were subsequently used to investigate performance patterns in terms of these domains.

Low, intermediate and high performers were defined relative to other students in the same year cohort. For each domain, each cohort was divided into three performance levels: low, intermediate and high. Low performers were students scoring below the 20$^{th}$ percentile; intermediate performers were students scoring between the 20$^{th}$ and 80$^{th}$ percentile; high performers were students scoring above the 80$^{th}$ percentile. Nine intersecting subdivisions of the study sample were thus identified for each year cohort: low, intermediate and high in *languages* (referred to as *lowL*, *intL* and *highL*), low, intermediate and high in *social studies* (referred to as *lowSS*, *intSS* and *highSS*), and low, intermediate and high in *science and math* (referred to as *lowSM*, *intSM* and *highSM*). Note that 109 (6.8%) students were low performers in all three domains, 507 (31.8%) were intermediate performers in all domains and 138 (8.6%) were high performers in all domains.

As noted, the approach taken to address the main research issues was to explore patterns in the relationships between domains at different performance levels and for different year cohorts. This was done in two ways. First, Pearson correlations were calculated between domains for each of the subdivisions described and correlation strength assessed using Cohen's (1992) guidelines. Correlation information was then used as a heuristic for detecting general patterns in the relationships between domains, which were subsequently subjected to significance testing. The Hotelling-Williams' test was used to compare dependent correlations (i.e., within a cohort) and Fisher's *z*-test (following transformation of the correlation coefficients) was used to compare correlations from independent samples (i.e., across cohorts) (Krishnamoorthy & Xia, 2007). Second, year-related differences in performance in each domain were explored for each performance level group. This was done using analyses of variance with domain scores as the outcome variables and year as between-subjects factor with post-hoc tests (Bonferroni correction).

## RESULTS

*Models of school performance*

Estimates of model-fit for the examined models are presented in Table 6.2. None of the unidimensional models achieved adequate fit on the combination of indices; thus, a single cognitive trait does not underlie outcomes in all content areas. For Year 1, the model-fit of the three models (B, C and D) in which language of schooling is related to languages and also to social studies or/and science and math was acceptable and

Table 6.2

*Estimates of CFA model-fit*

| Model | $\chi^2$ | df | *p* | SRMR | CFI | RMSEA | | | CAIC |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | RMSEA | 90% CI | *p-close*[e] | |
| Year 1 | | | | | | | | | |
| Unidimensional | 125.07 | 14 | <.001 | .05 | .93 | .14 | .12 - .16 | <.001 | 223.33 |
| Model A[a] | 57.68 | 11 | <.001 | .03 | .97 | .10 | .08 - .13 | <.001 | 176.99 |
| Model B[b] | 25.02 | 10 | .005 | .02 | .99 | .06 | .03 - .09 | .25 | 151.35 |
| Model C[c] | 23.72 | 10 | .008 | .02 | .99 | .06 | .03 - .09 | .30 | 150.06 |
| Model D[d] | 21.16 | 9 | .01 | .02 | .99 | .06 | .03 - .09 | .31 | 154.51 |
| Year 2 | | | | | | | | | |
| Unidimensional | 455.17 | 27 | <.001 | .07 | .87 | .16 | .15 - .18 | <.001 | 588.35 |
| Model A[a] | 128.40 | 24 | <.001 | .04 | .97 | .08 | .07 - .10 | <.001 | 283.76 |
| Model B[b] | 123.38 | 23 | <.001 | .03 | .97 | .08 | .07 - .10 | <.001 | 286.15 |
| Model C[c] | 116.22 | 23 | <.001 | .03 | .97 | .08 | .07 - .10 | <.001 | 278.99 |
| Model D[d] | 115.81 | 22 | <.001 | .03 | .97 | .08 | .07 - .10 | <.001 | 285.98 |
| Year 3 | | | | | | | | | |
| Unidimensional | 582.27 | 27 | <.001 | .08 | .81 | .19 | .17 - .20 | <.001 | 714.92 |
| Model A[a] | 92.80 | 24 | <.001 | .04 | .98 | .07 | .05 - .08 | .01 | 247.56 |
| Model B[b] | 82.56 | 23 | <.001 | .03 | .98 | .07 | .05 - .08 | .04 | 244.70 |
| Model C[c] | 85.45 | 23 | <.001 | .03 | .98 | .07 | .05 - .08 | .03 | 247.59 |
| Model D[d] | 82.23 | 22 | <.001 | .03 | .98 | .07 | .05 - .08 | .03 | 251.74 |

Notes. [a] 3-factor model in which language of schooling is related only to the *languages* domain. [b] 3-factor model in which language of schooling is related to *languages* and *social studies* domains. [c] 3-factor model in which language of schooling is related to *languages* and *science and math* domains. [d] 3-factor model in which language of schooling is related to all domains. [e] *p-close* tests the null hypothesis of close fit.

equivalent (SRMR = .02; RMSEA = .06; CFI = .99). The CAIC favoured Model B (i.e., language of schooling related to both languages and social studies; CAIC = 151.35) and Model C (i.e., language of schooling related to both languages and science and math; CAIC = 150.06), which differed only marginally. However, in Model C, the standardised loading of language of schooling on the languages domain became very low and non-significant ($\beta = .12$, $p = .23$). Model C thus seemed to reflect an alternative view of the nature of both domains that would not support a meaningful comparison of performance across years. For this reason, Model B was selected for the further analyses. Nonetheless, an interpretation of Model C is included in the Discussion.

For Year 2, the model-fit of all the 3-factor models was acceptable (SRMR ≤ .04; RMSEA = .08; CFI = .97). The CAIC preferred Model A (i.e., language of schooling related only to the languages domain; CAIC = 283.76) and Model C (i.e., language of schooling related to both languages and science and math; CAIC = 278.99). However, in Model C, the standardised loading of language of schooling on the science and math domain ($\beta = .19$) was too low to be of relevance (Kline, 2005), so Model A was selected for the further analyses.

For Year 3, the model-fit of all the 3-factor models was acceptable (SRMR ≤ .04; RMSEA = .07; CFI = .98). Model B (i.e., language of schooling related to both languages and social studies) had the lowest CAIC (244.70). However, the standardised loading of language of schooling on the social studies domain ($\beta = .21$) was too low to be of relevance (Kline, 2005). The CAIC then favoured Models A (i.e., language of schooling related only to the languages domain; CAIC = 247.56) and Model C (i.e., language of schooling related to both languages and science and math; CAIC = 247.59), which differed only marginally. In Model C, the standardised loading of language of schooling on science and math ($\beta = .12$) was again too low to be of relevance (Kline, 2005), so Model A was selected for the further analyses.

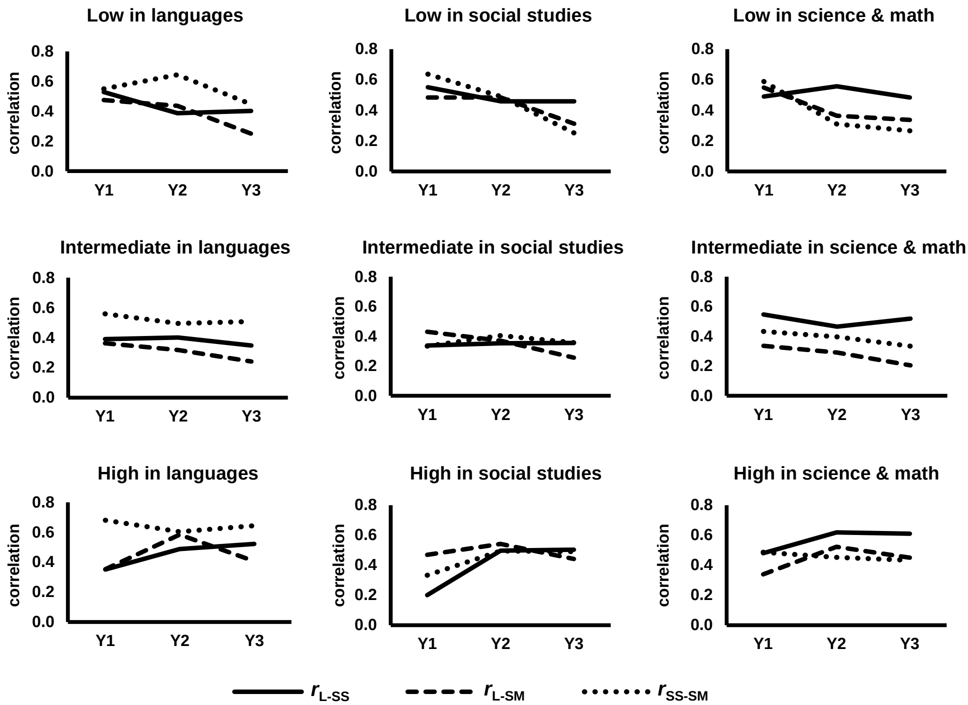
In short, for Year 1, proficiency in the language of schooling was substantially related to both the languages and the social studies domains. In Years 2 and 3, language of schooling was not substantially related to performance in social studies or science and math. Standardised and unstandardised estimates for the selected models are presented in Table 6.3.

Table 6.3
*Standardised and unstandardised estimates for selected CFA models*

| Indicator | Latent factor | β | *B* | *SE* | *p* | $R^2$ |
|---|---|---|---|---|---|---|
| | | | Year 1 (Model B) | | | |
| English | Languages | .72 | 1.00 | | | .52 |
| French | Languages | .90 | 1.32 | .09 | < .001 | .81 |
| Dutch | Languages | .37 | 0.41 | .07 | < .001 | .59 |
| Dutch | Social studies | .46 | 0.54 | .08 | < .001 | |
| Geography | Social studies | .85 | 1.00 | | | .73 |
| History | Social studies | .82 | 0.96 | .05 | < .001 | .68 |
| Biology | Science & math | .84 | 0.93 | .06 | < .001 | .70 |
| Math | Science & math | .75 | 1.00 | | | .56 |
| | | | Year 2 (Model A) | | | |
| English | Languages | .75 | 1.00 | | | .57 |
| French | Languages | .84 | 1.35 | .06 | < .001 | .71 |
| German | Languages | .84 | 1.17 | .06 | < .001 | .71 |
| Dutch | Languages | .78 | 0.84 | .04 | < .001 | .60 |
| Geography | Social studies | .92 | 1.00 | | | .85 |
| History | Social studies | .78 | 0.83 | .04 | < .001 | .60 |
| Biology | Science & math | .80 | 0.94 | .05 | < .001 | .64 |
| Physics | Science & math | .80 | 1.04 | .06 | < .001 | .64 |
| Math | Science & math | .71 | 1.00 | | | .50 |
| | | | Year 3 (Model A) | | | |
| English | Languages | .62 | 1.00 | | | .39 |
| French | Languages | .84 | 1.41 | .09 | < .001 | .71 |
| German | Languages | .80 | 1.28 | .08 | < .001 | .64 |
| Dutch | Languages | .73 | 0.83 | .06 | < .001 | .54 |
| Geography | Social studies | .87 | 1.00 | | | .76 |
| History | Social studies | .79 | 1.02 | .05 | < .001 | .62 |
| Chemistry | Science & math | .83 | 1.08 | .05 | < .001 | .68 |
| Physics | Science & math | .90 | 1.33 | .06 | < .001 | .81 |
| Math | Science & math | .76 | 1.00 | | | .57 |

*Performance patterns*

Figure 6.2 and Table 6.4 present the bivariate correlations between domains and across years.




Note. *r* is Pearson correlation, L is *languages*, SS is *social studies*, SM is *science and math*.

*Figure 6.2*. Graphs of bivariate domain correlations by domain, performance level and year

*Low performers.* There were moderate to strong positive and significant correlations between performance in all three domains in all years ($r = .25$ to .64). There appeared to be a robust and stable relationship between *languages* and *social studies* for all low performing groups (pattern 1). By comparison, relationships between *science and math* and the other domains appeared to decrease from Year 1 to Year 3 (pattern 2). Both patterns were tested with Fisher's *z*-tests (i.e., comparing correlations from independent samples). For pattern 1, *languages-social studies* coefficients were compared in Year 1 and Year 3. Results were non-significant for all groups, indicating that coefficients were similar across these years; pattern 1 was therefore confirmed. For pattern 2, *languages-science and math* and *social studies-science and math* coefficients were compared in Year 1 and Year 3 with one-tailed tests to reflect the hypothesised direction (i.e., Year 1 > Year 3). Regarding the

Table 6.4

*Bivariate domain correlations by domain, performance level and year*

| | Year 1 | | | | Year 2 | | | | Year 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $N$ | $r_{L\text{-}SS}$ | $r_{L\text{-}SM}$ | $r_{SS\text{-}SM}$ | $N$ | $r_{L\text{-}SS}$ | $r_{L\text{-}SM}$ | $r_{SS\text{-}SM}$ | $N$ | $r_{L\text{-}SS}$ | $r_{L\text{-}SM}$ | $r_{SS\text{-}SM}$ |
| Low in: | | | | | | | | | | | | |
| Languages | 82 | .53** | .47** | .55** | 120 | .39** | .44** | .64** | 119 | .40** | .25** | .45** |
| Social studies | 82 | .55** | .48** | .64** | 117 | .46** | .48** | .49** | 117 | .46** | .31** | .25** |
| Science & math | 78 | .49** | .55** | .59** | 122 | .56** | .36** | .31** | 116 | .48** | .34** | .26** |
| Intermediate in: | | | | | | | | | | | | |
| Languages | 247 | .39** | .36** | .56** | 361 | .40** | .32** | .50** | 348 | .35** | .24** | .51** |
| Social studies | 247 | .34** | .43** | .33** | 359 | .35** | .37** | .41** | 350 | .36** | .26** | .36** |
| Science & math | 248 | .55** | .34** | .43** | 359 | .47** | .29** | .40** | 351 | .52** | .21** | .33** |
| High in: | | | | | | | | | | | | |
| Languages | 82 | .35** | .35** | .68** | 120 | .49** | .58** | .60** | 117 | .52** | .41** | .64** |
| Social studies | 82 | .20 | .47** | .33** | 125 | .50** | .54** | .50** | 117 | .50** | .44** | .49** |
| Science & math | 85 | .48** | .34** | .49** | 120 | .62** | .52** | .45** | 117 | .61** | .45** | .43** |

Notes. $r_{L\text{-}SS}$: Pearson correlation between *languages* and *social studies*. $r_{L\text{-}SM}$: Pearson correlation between *languages* and *science and math*. $r_{SS\text{-}SM}$: Pearson correlation between *social studies* and *science and math*. **Correlation is significant at the 0.01 level (2-tailed).

relationship between *languages* and *science and math*, results were significant for the *lowL* and *lowSM* groups ($z_{Y1>Y3} = 1.78$, $p = .04$ for both groups) and showed a tendency in this direction for the *lowSS* group ($z_{Y1>Y3} = 1.40$, $p = .08$). For the relationship between *social studies* and *science and math*, results were significant for the *lowSS* and *lowSM* groups ($z_{Y1>Y3} = 3.38$, $p < .001$ and $z_{Y1>Y3} = 2.71$, $p = .003$ respectively) but not for the *lowL* group. Thus, pattern 2 was partially but not fully confirmed. Modest year-related differences in domain scores were found for all low performers in domains outside the defined area of weakness. For the *lowL* group, performance was higher in Year 3 than in Year 1 in *science and math* ($F(2,318) = 3.97$, $p = .02$, $\eta_p^2 = .02$; $p_{Y3>Y1} = .02$). For the *lowSS* and *lowSM* groups, *languages* performance was higher in Year 3 than in Year 1 ($F(2,313) = 3.26$, $p = .04$, $\eta_p^2 = .02$; $p_{Y3>Y1} = .03$ and ($F(2,313) = 5.28$, $p = .006$, $\eta_p^2 = .03$; $p_{Y3>Y1} = .005$ respectively).

*Intermediate performers*. There were moderate to strong positive and significant correlations between performance in all three domains in all years ($r = .21$ to $.56$). Relationships between *languages* and *social studies* and *social studies* and *science and math* appeared to be stable for all intermediate groups (pattern 3), while the relationship between *languages* and *science and math* appeared to decrease from Year 1 to Year 3 (pattern 4). Both patterns were tested with Fisher's $z$-tests. For pattern 3, *languages-social studies* and *social studies-science and math* coefficients were compared in Year 1 and Year 3. Results were non-significant for all groups, indicating that coefficients were similar across these years; thus, pattern 3 was confirmed. For pattern 4, *languages-science and math* coefficients were compared in Year 1 and Year 3 with one-tailed tests to reflect the hypothesised direction (i.e., Year 1 > Year 3). Results were significant for the *intSS* and *intSM* groups ($z_{Y1>Y3} = 2.39$, $p = .01$ and $z_{Y1>Y3} = 1.69$, $p = .046$ respectively) and showed a strong tendency in this direction for the *intL* group ($z_{Y1>Y3} = 1.61$, $p = .054$). Thus, pattern 4 was largely confirmed. No year-related differences in domain scores were found by the ANOVAs. This means that relative performance levels for these students did not differ across year cohorts.

*High performers*. There were medium to strong positive and significant correlations between performance in all three domains in all years ($r = .33$ to $.68$), with one exception ($r = .20$, n.s.). The relationship between *languages* and *social studies* was particularly strong for all high performing groups in Year 3 ($r = .50$ to $.61$) and appeared to increase from Year 1 to Year 3 (pattern 5). This was tested with one-tailed Fisher's $z$-tests comparing *languages-social studies* coefficients in Year 1 and Year 3. Results were significant for *highSS* students ($z_{Y3>Y1} = 2.40$, $p = .01$) and showed a tendency in this direction for the *highL* group ($z_{Y3>Y1} = 1.45$, $p = .07$). Pattern 5 was therefore only partially confirmed. Regarding year-related differences in domain scores, *social studies* performance of the *highSS* group was higher in Year 3 than in Year 1 ($F(2,321) = 3.16$, $p = .04$, $\eta_p^2 = .02$; $p_{Y3>Y1} = .06$) but there were no year-related differences otherwise.

Finally, at all *science and math* performance levels, *science and math* appeared to be least related to the other domains by Year 3 (pattern 6). This was tested with the Hotelling-Williams' test (i.e., comparing dependent correlations within a cohort). Year 3 *science and math* coefficients ($r_{L\text{-}SM}$ and $r_{SS\text{-}SM}$) were compared to the *languages-social studies* coefficient ($r_{L\text{-}SS}$), using one-sided tests to reflect the direction of the hypothesis (i.e., $r_{L\text{-}SS} > r_{L\text{-}SM}$ and $r_{L\text{-}SS} > r_{SS\text{-}SM}$). In the *lowSM* group, the *social studies-science and math* relationship was significantly weaker than the *languages-social studies* relationship ($T_2(r_{L\text{-}SS} > r_{SS\text{-}SM}) = 2.28$, $p = .01$), while the *languages-science and math* relationship showed a tendency to be weaker than the *languages-social studies* relationship ($T_2(r_{L\text{-}SS} > r_{L\text{-}SM}) = 1.48$, $p = .07$). Both results were significant for the *intSM* group ($T_2(r_{L\text{-}SS} > r_{L\text{-}SM}) = 5.82$, $p < .001$; $T_2(r_{L\text{-}SS} > r_{SS\text{-}SM}) = 3.20$, $p = .001$) and the *highSM* group ($T_2(r_{L\text{-}SS} > r_{L\text{-}SM}) = 2.04$, $p = .02$; $T_2(r_{L\text{-}SS} > r_{SS\text{-}SM}) = 2.28$, $p = .01$). Thus, pattern 6 was largely confirmed.

## DISCUSSION

This large-scale, cross-sectional study investigated the performance profiles of students in academic tracks in the first three years of secondary school within a stratified educational system. The approach taken was to examine and compare patterns of performance across academic domains and year cohorts for low, intermediate and high performers. This could shed light on individual and educational factors differentiating low performers from better performing students that may be amenable to intervention. With these insights, it may be possible to help low-performing students in academic tracks to improve their performance and keep on track.

Preliminary to the main analyses, confirmatory factor analysis (CFA) of school grades established that school performance in the study sample reflected three latent factors corresponding to the commonly-delimited domains of *languages* (i.e., language of schooling and modern foreign languages), *social studies* (i.e., geography and history) and *science and math* (i.e., the natural sciences and mathematics). For the first year of secondary school, it was found that proficiency in the language of schooling was related to both the *languages* and *social studies* domains. However, proficiency in the language of schooling was no longer substantially related to the *social studies* or *science and math* domains after the first year. This was expected on the basis of research into vocabulary development (Benelli et al., 2006; Nippold et al., 1999), which suggested a higher role of proficiency in the language of schooling for younger adolescents.

The CFA results were used to calculate factor quotient scores, which represented performance in each domain. Subsequently, low, intermediate and high performers were identified per domain and year cohort. Performance levels across domains were substantially and positively correlated for all groups. This means that, in general, students who do well in one domain also do well in others, and that students who do less well in one domain also do less well in others. Ostensibly, this appears to support a stratified educational approach whereby students are given instruction in all content areas at a level corresponding to their presumed general ability. However, the study results provide more nuanced insights into this issue. For a start, the positive relationship between domains cannot be explained by some general factor (like general intelligence) underlying performance in all domains, as unidimensional models were not supported by the CFA. This finding argues against stratification according to some conceptualisation of general ability. Second, at all *science and math* performance levels, *science and math* was least related to other content areas in Year 3. Furthermore, relationships between *science and math* and both other domains decreased from Year 1 to Year 3 for most low performers, while the relationship between *science and math* and *languages* decreased from Year 1 to Year 3 for intermediate performers. This suggests that *science and math* competence may develop along a different path to competence in other areas and, consequently, that instruction in these areas may not need to be pitched at the same educational level as other content areas.

Taken together, these findings argue for more modular forms of schooling (as opposed to singular tracks) in which students are offered instruction at the level appropriate to their competence in any particular domain or even individual content area. In support of this argument, examination of year-related differences in domain scores found that relative performance of low performers improved to some degree in domains outside the defined area of weakness. Again, this suggests that performance is not an across-the-board phenomenon and that students may be receptive to different levels of instruction in different domains.

Nonetheless, the positive relationship between competence in different domains remains to be explained. The results are compatible with the notion that content area understanding may interact in a form of co-construction or bootstrapping. In other words, building competence in one area may facilitate understanding and proficiency in other areas. This effect was seen in high performers, with high competence in *languages* and *social studies* appearing to interact in a beneficial - and for many of these students increasingly beneficial - way. In low and intermediate performers also, the positive relationship between *languages* and *social studies* remained robust throughout the years studied. This suggests that general language skills may play a key role in performance in other academic domains - particularly *social studies* - and that low language skills may underpin lower performance in these areas (Gomez & Gomez, 2007; Moje, 2002).

It is important to realise that these results pertain to the *languages* domain as a whole, which includes not only the language of schooling but also modern foreign languages. Indeed, research indicates that foreign language learning is beneficial to cognitive development in several ways. For example, attention, memorisation, analysis of language input and construction of meaning are all stimulated by foreign language learning while also being valuable for academic activity in general (Baker, 2006; Robinson, 2008). Thus, the relationship between performance in the *languages* domain and other academic areas may be mediated by these kinds of general skills.

The different findings regarding the role of language in school performance can be given some extra clarification, if one assumes that performance measures are a reasonable reflection of learning and understanding. On the one hand, there are indications that general language skills are important for content area understanding, and that this is the case throughout the first years of secondary school. On the other hand, proficiency specifically in the language of schooling appears not to be strongly related to *social studies* and *science and math* understanding after the first year.

Our interpretation of these results is, first, that general skills that characterise the *languages* domain are highly relevant and transferable academic skills required for progression in any domain. They can therefore be expected to remain of importance throughout learning, as the results show. Second, both the model in which language of schooling was an indicator for *social studies* (Model B) and the model in which language of schooling was an indicator for *science and math* (Model C) fit the data well for the first year. Thus, *initial* access to many content areas may depend specifically on proficiency in the language of schooling. Indeed, first year textbooks contain a myriad of abstract concepts and terms that students meet for the first time, such as 'prosperity', 'reformation', 'modernisation', 'social segregation', 'mitosis', 'ecosystem', 'exponential', etc. Given the research on vocabulary acquisition described earlier, it seems reasonable that students who are more proficient in using and understanding more abstract vocabulary in the language of schooling would be better able to access content areas through these kinds of texts. However, it is likely that once initial content knowledge has been acquired, it can be used to scaffold the acquisition and integration of new domain-specific information (Gagné, 1977; Ivie, 1998), even if language proficiency is not optimal. In other words, acquisition of this knowledge could mitigate initial deficiencies in the language of schooling. Alternatively, information gap theory suggests that a small increase in knowledge through initial learning can increase students' curiosity for acquiring new knowledge and motivation to learn (Loewenstein, 1994). In short, initial reliance on proficiency in the language of schooling may fade once students begin to acquire domain expertise.

*Educational implications*

Taken together, the results of this study provide important implications for educational research and practice. In particular, the importance of general language skills supports the current educational emphasis on language improvement. For example, the adolescent literacy movement sees both general and specific language skills as going hand-in-hand with content learning, and much effort is put into developing and implementing educational methods that integrate these skills (see Pearson, Moye, & Greenleaf, 2010, for an overview). While it is essential to develop general language skills to improve students' academic prospects - and we do not advocate 'teaching around the text' as a universal maxim - if only *initial* content learning depends strongly on proficiency specifically in the language of schooling, this opens up new avenues to be explored.

Current educational methods rely heavily on textual materials for knowledge acquisition, storage and interpretation. This unavoidably leads to reliance on language skills for learning. Clearly, this disadvantages groups whose language skills (general or specific) may be inadequate to deal with this method of content delivery. At risk in this respect are students whose native language is different from the language of schooling, and students from environments with little language stimulation. To this end, curriculum developers and educational publishers could develop alternative methods of instruction to help these students acquire initial knowledge of content areas. As this study indicates, once initial knowledge has been acquired, proficiency in the language of schooling may become less essential for further acquisition of content knowledge. In this way, it may be possible to prevent language deficiencies from retarding progression in content area understanding. Indeed, research has shown that individuals with low verbal comprehension skills can outperform more skilled readers on comprehension measures when they have more knowledge of the subject matter treated (Zwaan & Radvansky, 1998). Furthermore, growing content area expertise may also be pivotal in bootstrapping language skills by providing a cognitive framework within which verbal information can be integrated to infer meaning (Gomez & Gomez, 2007; Zwaan & Radvansky, 1998).

Alternative methods of instruction could be developed based on visuospatial materials. Visuospatial reasoning can provide an approach to learning and problem-solving that can supplement or even replace logical or analytical thinking, for example by using diagrams or drawings to solve problems or conceptualise functions (Ramadas, 2009). Moreover, visuospatial abilities correspond highly with students' abilities to process information when dealing with mathematical, science or geoscience content areas (Kozhevnikov, Motes, & Hegarty, 2007; Liben & Titus, 2012; Sanchez, 2012; Wu & Shah, 2004). From this perspective, educators should be encouraged to pay attention to visuospatial as well as language skills. This could mean providing training in visuospatial skills (e.g., Sanchez, 2012) or supplementing textual materials with visuospatial tools and materials (for example, Geographic Information Systems, which use innovative technology to process and present spatial information (Favier & Van der Schee, 2009)). Intriguingly, there is evidence that certain language features - namely the breadth of semantic associations (such as used in forming abstract categories) - may be closely tied to visuospatial thinking (Brugger, Loetscher, Graves, & Knoch, 2007). In that case, it could be hypothesised that improving visuospatial skills could also benefit language development.

The finding that building competence in one domain may facilitate understanding and proficiency in other domains also has implications for educational practice. Rather than segregating individual content areas, schools could integrate multiple areas in project-based learning. Project-based learning is designed to expand views of subject matter, to promote links between diverse subject matter and to develop higher-order thinking processes (Cochrane, 2004; Curtis, 2002). Project-based learning has also been used successfully to improve school performance, self-image and learning motivation of low-achieving secondary school students (Doppelt, 2003).

Finally, the findings argue for a revised approach to early stratification in secondary education and are therefore particularly relevant in light of ongoing debate as to the advantages and disadvantages of early tracking systems. On the one hand, early tracking allows vocational training to start earlier, thereby equipping students with specific skills which can enhance their job productivity on leaving school (Andersen & Van de Werfhorst, 2010; Shavit & Müller, 2000; Wolbers, 2007). On the other hand, national and international studies show that early tracking perpetuates and even magnifies social inequalities in educational outcomes (lower-class students typically being placed in lower tracks) and reduces average achievement (Hanushek & Wößmann, 2006; Schütz, Ursprung, & Wößmann, 2008; Shavit & Müller, 2000; Van de Werfhorst & Mijs, 2010). The findings of the present study support a more comprehensive approach - in any case during the first three years of secondary school - whereby instruction is uncoupled from the notion of general ability and becomes more modular with respect to intra-individual differences in ability. In this way, students could benefit from domain-specific instruction that is more specifically geared to their own level of competence in that domain and intra-individual differences could be better served.

*Limitations and future research*

Certain limitations of and issues raised by this study could be addressed in future research. First, the approach and methods should be replicated with other samples from stratified educational systems to establish the extent to which the findings are robust across various grading contexts and systems. This recommendation is especially pertinent in light of known issues described earlier that can affect performance measurement. Second, it is possible that other structural models than those examined here would fit the data better; future research could specifically examine how best to model the underlying dimensions of school performance. Third, school grades were transformed for analysis using the Deus method (see Appendix) to correct for structural differences in levels of test difficulty between tracks. Although we consider this method to give a reasonable approximation of the required corrections, we would welcome research that investigates alternative methods for cross-track performance comparison in the Dutch situation. Fourth, the examination of year-related differences was based on cross-sectional data, which tempers conclusions drawn with respect to developmental aspects. The study was also not designed to address causal issues, such as the effects of being placed in certain tracks. Longitudinal studies and national and international studies into tracking effects such as discussed above could explore both these aspects.

Finally, results were not differentiated with respect to sex, though it is conceivable that the relationships reported here could differ for boys and girls. It is notable that boys are overrepresented in lower educational tracks, repeating a year, switching to lower tracks and school drop-out (Driessen & Van Langen, 2010). There are reasons to believe that these effects could arise from slight differences in the dynamics of brain development that disfavour boys with respect to certain types of educational materials, at least over the age-period investigated in this study. Specifically, research indicates that boys and girls differ with respect to the development of certain cortical structures and functions that are essential to learning. In particular, language-related brain areas are found to be larger and to mature earlier in adolescent females, while areas engaged in certain visuospatial tasks are larger and mature earlier in males (Giedd, Raznahan, Mills, & Lenroot, 2012). Although structural brain differences do not in themselves say anything about cognition and behaviour, they may be further strengthened by socialisation differences in experience, opportunities, expectations etcetera (Halpern, 2012). In consequence, in early to mid adolescence - namely during the period covered by this study - girls may be at advantage with respect to language-related aspects of their schooling such as processing textual materials, while boys may be at advantage with respect to visuospatial tasks and materials (Clements-Stephens, Rimrodt, & Cutting, 2009). Given that most secondary school educational methods rely heavily on textual rather than visuospatial materials, as noted, it is conceivable that girls' superiority with language, which also shows consistently in behavioural studies (e.g., Driessen & Van Langen, 2010) and PISA results (OECD, 2010; Stoet & Geary, 2013), contributes to both girls' relatively greater school success and boys' overrepresentation among the lower performers. It would be of interest to investigate this issue in a subsequent study.

To conclude, this study provides new insights into patterns of school performance in academic tracks within a stratified educational system, with a particular focus on low performers. The study presents a nuanced perspective on the role of certain factors - notably general and specific language skills - in shaping individual learning outcomes. These insights could help educators and curriculum developers to accommodate individual differences and developmental changes, and to develop learning materials that may help low performers to keep on track.

**Acknowledgements**

The authors thank Paul Kirschner, Annemarie Boschloo, Wim van der Elst, Renate de Groot and Aukje Aben for their contributions to this study. We also thank two anonymous reviewers for their comments on an earlier version of this article.

**Appendix: Correction for structural track grading differences**

First, it must be noted that there is no fixed method for equating school grades across tracks in the Netherlands, and that cross-track comparison is rare. In the most general terms, a difference of approximately 1 point in grade value denotes the difference in level of test difficulty between adjacent tracks. We considered an appropriate method for correcting for these structural differences to be the Deus method developed by the Dutch National Board of Examinations (CvE) as part of their mandate from the Dutch Government to establish examination norms and guarantee the level and degree of difficulty of national exams. The Deus method has been implemented since 2000 by the Dutch National Institute for Educational Measurement (CITO) to derive national examination grades from test scores (Sanders & Verstralen, 2011). The method employs a norm variable (the so-called 'N-term') that in practice takes a value between 0 and 2 to correct for varying levels of test difficulty. The value 1 indicates that the test is at the desired level of difficulty, while values below 1 and above 1 are used to correct for levels of difficulty that are too low and too high respectively.

In the present study, the pre-university and combined tracks ($N = 1066$) were used as the reference level (note that end-of-year grades for the combined track in the participating schools were at pre-university level) with an N-term of 1. In this situation, tests in the higher general secondary track ($N = 530$) can be considered to be at a too low level of difficulty, and should therefore be corrected with an N-term below 1. In the most generally representative situation (i.e., 1 point difference in grade value across adjacent tracks), students in both tracks should have approximately the same raw grade average, while corrected track grade averages should differ by 1 point after transformation (i.e., students in the lower track should have a corrected grade average of approximately 1 less than students in the higher track). The N-terms used to correct grades in the higher general secondary track were therefore calibrated (with values between 0 and 1) per school and year cohort, so that the corrected grades approached the most representative situation.

The Deus transformation uses the formula: $C = (9/L)*S + N$, where $C$ represents the final grade, $L$ is the maximum test score, $S$ is the individual's test score and $N$ is the N-term (note that there are additional restrictions to ensure that the transformed grade does not go beyond the grading scale extremities). In the present case, we considered the original school grades as representing individuals' test scores ($S$), with a maximum of 10 ($L$). Thus, grades for the pre-university and combined tracks were transformed with the formula $C = (9/10)*S + 1$ and grades for the higher general secondary track were transformed with the formula $C = (9/10)*S + \text{calibrated N-term } [\text{range } 0,1]$.