Journal for Educational Research Online
Journal für Bildungsforschung Online
Volume 11 (2019), No. 1, 166–189

Olaf Köller, Jennifer Meyer, Steffani Saß & Jürgen Baumert

New analyses of an old topic: Effects of intelligence and motivation on academic achievement

Abstract

This study focuses on a topic with a long tradition in educational psychology. In a large data set with several achievement measures we investigated the effects of intelligence and motivation on academic achievement in three domains, namely, German, mathematics, and English, using three different achievement measures (standardized tests, grades, and final written exams) in a sample of upper secondary students (N = 3 775; Grade 13; 54.8 % female; age M = 19.92 years) in Germany. Furthermore, we focused on grade point average (GPA) as a general achievement indicator at the end of upper secondary school. First, we aimed to replicate previous results on the predictive power of intelligence and motivation for achievement. Second, we aimed to extend the large body of existing research by adding final written exams – school-based performance tests – as an additional measure. Our findings indicate that motivation had stronger effects on achievement than intelligence did. This was particularly true for the domain-specific achievement measures. Motivation had the strongest effects on grades, followed by final exams. The effects of intelligence were comparatively stronger for standardized achievement tests. Overall, the findings suggest that both intelligence and motivation are important predictors of achievement and that this is true for all kinds of achievement measures.

Keywords

Intelligence; Motivation; Domain-specific achievement; Upper secondary education; Achievement measures

Olaf Köller (corresponding author) · Jennifer Meyer · Steffani Saß, Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, 24118 Kiel, Germany
e-mail: koeller@ipn.uni-kiel.de
jmeyer@ipn.uni-kiel.de
sass@ipn.uni-kiel.de

Jürgen Baumert, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany
e-mail: sekbaumert@mpib-berlin.mpg.de

Neue Analysen zu einem alten Thema: Effekte von Intelligenz und Motivation auf Schulleistungen

Zusammenfassung

Die vorliegende Arbeit behandelt ein Thema mit langer Tradition in der pädagogischen Psychologie. Wir nutzen einen großen Datensatz mit verschiedenen Leistungsmaßen (Zeugnisnoten, standardisierte Leistungstests und schriftliche Abiturprüfungen), um die Effekte von Intelligenz und Motivation auf Schulleistungen in drei Fächern (Deutsch, Mathematik und Englisch) zu untersuchen. Zusätzlich werden Effekte auf die Abiturdurchschnittsnote analysiert. Die Stichprobe besteht aus 3 775 Schülerinnen und Schülern (54.8 % weiblich; mittleres Alter M = 19.92 Jahre) am Ende der 13. Jahrgangsstufe in deutschen Gymnasien. In einem ersten Schritt werden frühere Befunde zur prädiktiven Kraft von Intelligenz und Motivation für Schulleistungen (Tests und Noten) repliziert. Darüber hinaus werden die umfangreichen früheren Befunde dahingehend ergänzt, dass wir schriftliche Abiturleistungen analysieren. Insgesamt zeigen unsere Auswertungen deutlich stärkere Effekte von motivationalen Maßen im Vergleich zu Intelligenzindikatoren. Die stärksten motivationalen Effekte zeigten sich in fachspezifischen Analysen auf Zeugnisnoten, gefolgt von zentralen Abschlussprüfungen. Die Effekte von Intelligenz sind bei standardisierten Schulleistungstests höher. Insgesamt belegen die Befunde aber die Relevanz beider Prädiktoren für die Erklärung von Schulleistungen, und zwar für alle Leistungsindikatoren.

Schlagworte

Intelligenz; Motivation; fachspezifische Schulleistungen; gymnasiale Oberstufe; Schulleistungsmaße

1. Introduction

There is an ongoing debate on the predictive power of intelligence and motivation for academic achievement and learning (for recent studies, see Jansen, Lüdtke, & Schroeders, 2016; Lotz, Schneider, & Sparfeldt, 2018). Despite the fact that both characteristics have been shown to predict academic outcomes, research is still examining the relative contribution of each characteristic regarding its influence on achievement and learning. Previous research has revealed that the type of achievement measure probably moderates the effect sizes of motivation and intelligence, that is, while motivational factors often strongly predict school grades, intelligence seems more powerful in predicting standardized test scores (Helmke, 1992; Lotz et al., 2018; Spinath, Spinath, Harlaar, & Plomin, 2006; Steinmayr & Spinath, 2009). Furthermore, some evidence suggests that the effects of intelligence and motivation differ depending on the domain (e.g., particularly strong effects of intelligence on mathematics achievement; cf. Roth, Becker, Romeyke, Schäfer, Domnick, & Spinath, 2015) and on whether achievement indicators are domain-specific or more general, that is, *grade point average* (GPA) as a global measure of achievement (Poropat, 2009) versus subject- or domain-specific achievement scores. The work presented here aimed to address all these potential moderating effects when investigating the effects of motivation and intelligence on academic achievement. We based the statistical analyses on a large sample of upper secondary students in their final year of schooling in Germany (Leucht, Kampa, & Köller, 2016). In the following, we present previous research on the relationship between intelligence, motivation, and academic achievement.

1.1 Intelligence and achievement

Intelligence is one of the most important factors for explaining individual differences in educational achievement. Many cross-sectional studies have found substantial correlations between intelligence test scores and educational achievement (e.g., Benson, Kranzler & Floyd, 2016; Kaufman, Reynolds, Liu, Kaufman & McGrew, 2012). A recent meta-analysis by Roth et al. (2015) revealed a population correlation of $\rho = .54$ for intelligence and various school grades. The findings were robust (with small moderator effects) across different intelligence measures, different school subjects, different grade levels, and gender. In addition, the correlations between intelligence and standardized achievement test scores in large-scale assessments (LSA) of mathematics achievement ranged from $r = .38$ to $r = .72$ (Saß, Kampa, & Köller, 2017). Lotz et al. (2018) also reported coefficients ranging between .60 and .70 and referred to the study by Baumert, Lüdtke, Trautwein, and Brunner (2009) in which the correlation between intelligence and standardized test scores was even higher. It is generally expected that brighter students learn more easily, are better in dealing with new material, and easily transfer existing knowledge and skills to new learning situations. All of this is intrinsic to intelligence (Jensen, 1998).

Some research, however, also suggests that standardized tests and intelligence tests partly measure the same skill, particularly in mathematics. Saß et al. (2017) pointed out that inductive and deductive reasoning, which is part of intelligence, is also essential to understand mathematical relations, draw conclusions, and apply mathematical knowledge (for similar arguments see Kaufman et al., 2012). The authors thus argue that items in standardized mathematics tests often measure both mathematics achievement and intelligence.

Irrespective of construct validity, many studies have also confirmed the long-term predictive power of intelligence for educational achievement (e.g., Laidra, Pullmann, & Allik, 2007; Schneider & Niklas, 2017). For example, Deary, Strand, Smith, and Fernandes (2007) found that intelligence at age 11 influenced final school exams at age 16 in diverse subjects (e.g., languages, mathematics, sciences, arts, and music).

1.2 Motivation and achievement

The question as to whether motivation predicts students' academic behavior has a very long history in educational psychology. Based on pioneer work by Murray (1938) and Atkinson (1957), numerous studies have investigated the relationship between motivation and achievement. In particular, studies measuring motivation by means of questionnaires found stable positive associations between achievement motivation and academic achievement (cf. Fink, 1962; Gough & Fink, 1964; Schneider & Green, 1977). Subsequent studies used diverse theoretical approaches such as achievement goals (Elliot & McGregor, 2001), intrinsic motivation (Deci & Ryan, 2002), self-efficacy (Kriegbaum, Jansen, & Spinath, 2015), interest (Köller, Schnabel, & Baumert, 2001; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005), or academic self-concept (Guay, Ratelle, Roy, & Litalien, 2010; Lotz et al., 2018; Marsh et al., 2005) to study the relationship between motivational variables and academic achievement.

The *expectancy-value theory* (EVT; Eccles et al., 1983; Wigfield & Eccles, 2000) provides a powerful framework to better understand the role that domain-specific motivational variables play in predicting academic achievement and academic choices. Wigfield and Eccles (2000) argue that expectation of succeeding in a task and the value assigned to the task influence achievement-related behavior. Students who expect to be successful and who value a task highly spend more effort and time working on the task, experience less anxiety, and, in the end, have a higher probability of solving the task. Academic choices, for example, course selection, are also predicted by both components. Students prefer courses for which their self-perceived competence is high and which they value.

In empirical studies, domain-specific self-concept often serves as an indicator of the expectancy component and domain-specific interest represents the value component (cf. Marsh et al., 2005). Both self-concept and interest usually show substantial correlations with achievement. The size of this correlation is moderated by the type of achievement measure; correlations with grades are stronger than with standardized achievement tests. For example, in their meta-analysis, Möller, Pohlmann, Köller, and Marsh (2009) found a correlation of r = .50 between mathematics self-concept and grades, while the coefficient was r = .37 for test scores. For verbal self-concept, the coefficients were r = .40 and r = .34, respectively. Marsh, Trautwein, Lüdtke, Köller, and Baumert (2006) found a similar pattern in a large sample (more than 4,000 students from upper secondary schools).

The reciprocal effects model (REM; Marsh & Craven, 2006; Möller, Zimmermann, & Köller, 2014) posits that subsequent achievement is not only affected by prior self-concept, but that subsequent self-concept is also affected by prior achievement. Students' results in academic tests and exams and their grades provide information for social upward and downward comparisons with their classmates. Success often leads to downward comparisons, thereby enhancing academic self-concept, and failure provokes upward comparison, thereby lowering academic self-concept. Achievement is thus not only a result but also a cause of self-concept. Such a REM has also been applied to the relationship between domain-specific interest and achievement (e.g., Köller et al., 2001). The associations found, however, are usually weaker (about r = .30, Schiefele, Krapp, & Winteler, 1992) than for self-concept. In a recent study, Jansen et al. (2016) ran secondary analyses of a large German data set comprising the achievement and interest data of more than 39,000 9th-graders from German secondary schools. Five different domains (German, mathematics, biology, chemistry, and physics) were considered. Correlations between interest and grades varied between r = .27 (biology) and r = .38 (mathematics) with a high degree of domain-specificity; interest and achievement were substantially correlated within one domain but coefficients across domains were small. No clear picture was found for standardized achievement tests; in general, correlations were small or close to zero. Comparing the effects of self-concept and interest on achievement, Lotz et al. (2018) concluded that self-concept seems to be the stronger predictor of achievement and that the effect of interest on achievement equals zero when self-concept is controlled for. Note that, in this context, the correlation between self-concept and interest is usually very high (above .50; cf. Marsh et al., 2005) and authors argue that feeling competent (high academic self-concept) increases interest (e.g., Marsh et al., 2005). Interest, on the other hand, seems more important in the case of academic choices (see Köller et al., 2001, for course selection).

Whereas the EVT specifically describes the role that domain-specific motivational factors play in academic achievement, other research approaches have tried to find more general predictors of global achievement measures such as GPA. On the basis of self-concept research, Guay et al. (2010) carried out a repeated measurement study that examined the relationship between academic self-concept and academic achievement in terms of a cumulative achievement measure. The correlations between self-concept and achievement were r = .60 (time point 1) and r = .52 (time point 2). In line with research on domain-specific self-concept (*reciprocal-effects model* (REM); Marsh & Craven, 2006), it has been argued that this correlation is the result of reciprocal effects between global academic self-concept and global achievement.

In a totally different line of research, the relation between personality factors and global achievement measures has been investigated (see De Raad & Schouwenburg, 1996; Poropat, 2009; Stumm, Hell, & Chamorro-Premuzic, 2011). In particular, conscientiousness has consistently been shown to be associated with academic achievement beyond cognitive ability in secondary education, with effect sizes of ρ = .23 (Poropat, 2009). To understand why conscientiousness predicts academic achievement, it can be useful to consider its connection to traits and types of behavior that are known to be crucial for school performance. For example, conscientiousness is closely related to motivational variables such as self-discipline, ambition, persistence, diligence, dutifulness, and grit (Credé, Tynan, & Harms, 2017; Dumfart & Neubauer, 2016; Ivcevic & Brackett, 2014; Schmidt, Fleckenstein, Retelsdorf, Eskreis-Winkler, & Möller, 2017). Further, conscientiousness has been linked to learning behavior that results in good grades (see Credé & Kuncel, 2008;

Kling, Noftle, & Robins, 2013), such as self-regulated learning (Bidjerano & Dai, 2007), goal orientation (Sorić, Penezić, & Burić, 2017), systematic studying and methodical learning styles (e.g., Geisler-Brenstein, Schmeck, & Hetherington, 1996; Komarraju, Karau, Schmeck, & Avdic, 2011), as well as academic effort, which refers to an individual's care and persistence regarding school work (see Trautwein, Lüdtke, Roberts, Schnyder, & Niggli, 2009; Trautwein, Lüdtke, Schnyder, & Niggli, 2006). In summary, previous research indicates that students who are more attentive to their school assignments (i.e., high in conscientiousness) tend to perform better at school (Poropat, 2009).

1.3 Simultaneous effects of intelligence and motivation on achievement

Although theory and research suggest that intelligence and motivation have unique as well as common effects on achievement, a large body of research has tried to find out which of the two factors is more important, that is, which explains more variance in student achievement (see the recent studies by Jansen et al., 2016; Lotz et al., 2018). A relatively old but methodologically sound study was carried out by Helmke (1992). In a sample of lower secondary school students, he found that cognitive variables accounted for 38 % of the variance in a standardized mathematics test, whereas domain-specific self-concept only explained 32 % of the variance. The picture changed for mathematics grades, where the predictive power of self-concept (57 % explained variance) was much higher than for cognitive variables (only 20 %). A similar pattern was reported by Lotz et al. (2018), who additionally had domain-specific interest as a second motivational variable. Interest, however, did not show substantial effects on grades or on test results after controlling for self-concept and intelligence. Jansen et al. (2016) only used domain-specific interest as a motivational measure and also found a pattern of results that was quite similar to Helmke (1992). In a repeated measurement study, Kriegbaum et al. (2015) ran cross-sectional and longitudinal path analyses to predict mathematics achievement in the *Programme for International Student Achievement* (PISA). The authors used different motivational measures and compared their effects on achievement after controlling for intelligence (in cross-sectional models) and prior achievement (in longitudinal models). The strongest motivational predictor in the cross-sectional analyses was task-specific self-efficacy with exactly the same regression weight as intelligence ($\beta = .41$). When prior achievement was controlled for, the effect of intelligence ($\beta = .20$) was slightly stronger than that of self-efficacy ($\beta = .15$).

Concerning global achievement measures, a lot of studies have focused on the effects of personality traits such as conscientiousness on GPA above and beyond intelligence. In a large study, Laidra, Pullmann, and Allik (2007) used personality measures and intelligence to predict GPA in several grade levels. Their results clearly show that intelligence had effects in all grade levels ($\beta > .40$). Furthermore, even after controlling for intelligence, conscientiousness had a substantial and unique effect on GPA in secondary school (β = .21). In addition to this work, other studies have combined domain-specific and global personality measures and examined the extent to which they predict academic achievement (cf. Spinath, Freudenthaler, & Neubauer, 2010). They found that domain-specific achievement was more strongly influenced by domain-specific motivational measures than by personality traits (see also Marsh et al., 2006).

Taken together, existing empirical studies clearly suggest that both factors, intelligence and motivation, contribute to academic achievement. This relationship, however, is moderated by the type of achievement measure used and by the domain investigated.

2. The present study

Much research has already been conducted in an attempt to better understand the predictive power of motivation and intelligence for academic achievement. There is broad consensus that both factors uniquely contribute to achievement and that their relative power depends on the domain and the type of achievement measure. Our study contributes to the existing literature. In a first step, we examined the predictive power of intelligence and domain-specific motivation on academic achievement in English, mathematics, and German, contrasting three achievement indicators: standardized achievement tests, final written exams, and school grades. In a second step, we examined the extent to which the GPA at the end of upper secondary school was predicted by intelligence and global motivational and personality measures, that is, academic self-concept and conscientiousness. We decided to use global motivational measures because they represent more general motivational tendencies that influence achievement-related behavior across domains and can therefore be expected to be predictive of such a global measure as GPA, which covers achievement outcomes in all off the different school subjects in upper secondary school.

Our first research goal was to replicate previous findings in a large and rich data set of students at the end of upper secondary school. In contrast to previous studies, the sample was academically positively selected, consisting only of those students prepared for university studies. Our second research goal was to add final written exams as a further achievement measure when examining the effects of intelligence and motivation. Only a small number of studies up until now have investigated the effects of intelligence and motivation on final written exams. These exams at the end of upper secondary education in Germany have important consequences for students as they make up one third of the final GPA and, consequently, are important for college admission. Students' preparation before the exams is crucial and dependent on students' learning motivation and behavior. In our sample of Grade 13, the students had to retrieve content that had been learned over at least two school years. However, in contrast to grades as an accumulated measure over a longer period of time, the final exam is a performance situation lasting four to six hours for the individual student. It is strongly characterized by a certain amount of pressure as all the knowledge and competencies have to be retrieved on one particular occasion with little room for compensation in the case of failure. Therefore, these final examinations are of particular interest when investigating the effects of intelligence and motivation.

Overall, we tested six hypotheses in a sample of cognitively selected upper secondary students:

1) We expected to find stronger effects of intelligence on mathematics achievement than on English and German achievement.
2) In the case of motivational variables, we did not have any different expectations for the different domains. However, we expected to find large positive effects of motivation on all domain-specific achievement measures.
3) Regarding domain-specific motivational measures, we predicted that academic self-concept would clearly outperform academic interest in predicting achievement in all three domains.
4) We further investigated whether the relationships between intelligence, motivation, and achievement differed as a function of the achievement indicator (achievement test scores vs. report card grades vs. final written exams)? We predicted that intelligence would have the strongest positive effects on standardized test scores and would have moderate effects on grades and final written exams.
5) For motivation, we expected to find a different pattern, namely, large positive effects on grades and final written exams and moderate effects on test scores.
6) Finally, we predicted that intelligence, conscientiousness, and academic self-concept, as a global achievement measure, would have positive effects on the GPA.

3. Method

The present study is based on secondary analyses of the *LISA 6* (Lernergebnisse an beruflichen und allgemeinbildenden Gymnasien in Schleswig-Holstein [Educational Outcomes of Students from Vocational and Academic Upper Secondary Schools]) large-scale assessment study. The study was conducted in the German Federal State of Schleswig-Holstein (N = 3 775; 54.8 % female; see Leucht et al., 2016). In Schleswig-Holstein, there are two school tracks in upper secondary education (Grades 11-13): the vocational and the academic track. The academic track refers to the traditional *Gymnasium*, which provides general preuniversity education, whereas the vocational track focuses on more applied subjects such as technical and economics courses, in addition to compulsory education (e.g., languages and mathematics). In both tracks, students can obtain the general higher education entrance qualification (*Abitur*) after successful completion of classes and final exams at the end of Grade 13. All participants worked on achievement tests and student questionnaires.

3.1 Procedures

Whereas participation in the questionnaires was voluntary, participation in the achievement tests was mandatory for all students at the randomly drawn academic-track schools (*N* = 1433 students from 17 schools) and at all of the vocational-track schools (*N* = 2342 students from 27 schools). Before commencing the study, written informed consent was obtained from all the participants and their parents. The study was carried out in accordance with the ethical guidelines for research with human participants as proposed by the American Psychological Association (APA). All of the study materials and procedures were approved by the *Ministry of Education, Science and Cultural Affairs of the Federal State of Schleswig-Holstein*. Data collection was managed by the *International Association for the Evaluation of Educational Achievement (IEA)*. Experienced university students who were trained by the IEA administered both tests and questionnaires. Although all students worked on achievement tests, data from the questionnaires were available for only *N* = 2234 students.

3.2 Measures

3.2.1 Academic achievement

Three domain-specific measures plus one global measure of academic achievement were used in the present investigation.

Report card grades. Domain-specific end-of-school-year report card grades (in German, mathematics, and English) were collected via school administration lists for Grade 13. In Germany, report card grades at upper secondary school range from zero to 15 points, with higher values indicating higher achievement.

Standardized tests: Data from standardized tests were available for mathematics and English, but not for German. Mathematics achievement was assessed using a 20-item mathematics test from the *National Educational Panel Study* (NEPS), which was based on the literacy concept and designed in line with the German educational standards and with international frameworks used in PISA (for details see Neumann et al., 2013; Kampa, Köller, Schmidt, & Leucht, 2016). Following the NEPS framework, mathematical competencies can be described with two dimensions: a) content areas in the field of mathematics (quantity [4 items], change and relationship [6 items], space and shape [3 items], and data and chance [7 items]), and b) the cognitive component of mathematical competence, covering processes related to solving mathematical problems. Five cognitive processes are measured in the test: technical abilities and skills (9 items), modeling (1 item), mathematical problem solving (4 item), using representational forms (5 item), and mathematical communication (1 item). The results of the *LISA 6* study (see also Kampa et al., 2016) provided evidence for the validity of the test. Techniques from the item response theory (IRT; one-parameter model with a background model), implemented in the computer software ConQuest 3.0 (Adams, Wu, & Wilson, 2012) were applied to compute five plausible values (PVs). The PV reliability was .91.

English achievement was measured with two different tests. The first one with listening and reading comprehension exercises consisted of items from the *German National Assessment* (GNA; e.g., Stanat, Böhme, Schipolowski, & Haag, 2016). The test items were designed to monitor the implementation of educational standards in Germany (see Köller, Knigge, & Tesch, 2010; see also Leucht, Fleckenstein, & Köller, 2016) and therefore represent competencies based on curricula for English language classrooms. Three to four tasks consisting of different items were presented in four 15-minute blocks. Blocks were balanced in difficulty and rotated in eight different booklets to control for position effects and performance decline with test duration *(multimatrix design)*. The sufficient reliability and validity of the test has been shown in previous studies; results can be linked to similar standardized tests such as PISA (see Fleckenstein, Leucht, Pant & Köller, 2016). Five PVs were computed for each student to obtain reliable proficiency scores for both skills.

In addition, students' proficiency in English was assessed by a short paper-and-pencil version of the *Test of English as a Foreign Language* (TOEFL), which was developed for the *Institutional Testing Program* (ITP) of the *Educational Testing Service* (ETS) in Princeton (cf. Köller & Trautwein, 2004). The test consists of three sufficiently reliable (internal consistency above .80) subscales, Listening Comprehension (LC); Structure and Written Expression (SWE) and Vocabulary and Reading Comprehension (VRC) which are usually combined to form a global proficiency score. Five PVs were estimated for each subtest and for the global score.

To get a combined English test score from the GNA and the TOEFL, we applied principal component analysis to the five subscales and used the first component (covering more than 70 % of the variance with all loadings above .70) for all further analyses. Note that all other components had Eigenvalues < 1 so that all five subtests were sufficiently represented by the first component. Because we had five PVs per dimension, we computed five global scores that reached a reliability of .90.

Final written exams. We collected information on the grades received in final written exams (*Abitur*) in all three domains via school administration lists. These grades also range from zero to 15 points, with higher values indicating better performance. Centralized *Abitur* tasks were provided by the *Ministry of Education, Science and Cultural Affairs of the Federal State of Schleswig-Holstein* for both school types. In mathematics, competencies in different fields were captured, for example, calculus and geometry, with coherent superordinate assignments consisting of several subtasks (KMK, 2002a). In English and German, assignments consisted of text comprehension (e.g., fictional and nonfictional texts) combined with essay writing tasks (e.g., KMK, 2002b).

Students had to sit final exams in only two of the three subjects. Therefore, not all students took final exams in all three subjects. As a consequence, in English, final exam results were available for N = 2,984 students, in mathematics for N = 3,074 students, and in German for N = 3,051 students.

GPAs. Students' GPA scores were collected from school administration lists at the end of upper secondary school. Students' GPAs in German upper secondary schools are calculated on the basis of a weighted combination of score-card grades of the last two school years (66.7 %) and grades in final exams (33.3 %). Scores range between 1 (*excellent*) and 4 (*sufficient*).

3.2.2 Intelligence

General cognitive ability was assessed using the figural (25 items) and verbal (20 items) reasoning subscales of the cognitive ability test (KFT4-12 R+; Heller & Perleth, 2000). Both subscales are indicators of fluid intelligence. The reliabilities of the subscales, as provided in the manual according to the Kuder-Richardson Formula 20, were satisfactory, ranging from α = .68 (verbal) to .81 (figural). Five PVs were calculated (PV reliability of .79) to obtain total scores for the students.

3.2.3 Motivation

Measures of domain-specific motivation, general academic motivation, and personality were used.

Domain-specific measures: We included self-concept and interest measures for the three domains, English, mathematics, and German, based on the EVT. Self-concept was measured by means of four items that have been used in several previous studies (e.g., Möller et al., 2014) and have been shown to be reliable and valid. An example item is "I have always been good in English/mathematics/German". Students responded on a four-point scale ranging from 1 (*totally disagree*) to 4 (*totally agree*). Reliabilities (Cronbach's alpha) reached values of .83 (English), .84 (German), and .89 (mathematics).

Interest was measured by means of four items for English, four items for mathematics, and three items for German. These items had also been used in previous studies (e.g., Marsh et al., 2005) and showed good psychometric properties. An example item is "I would like to have more lessons in English/mathematics/German". Students again responded on a four-point scale ranging from 1 (*totally disagree*) to 4 (*totally agree*). Reliabilities (Cronbach's alpha) reached values of .77 (English), .82 (German), and .88 (mathematics).

Global measures: Four items, again with a four-point response format, were used to assess academic self-concept. The items were taken from Jopt (1978). An example is "I often think that I am not as bright as my classmates". The items reached an internal consistency of alpha = .83.

The four items on conscientiousness were taken from the German short version of the Big Five Inventory (BFI-K; Rammstedt & John, 2005). An example item is “I do tasks thoroughly”. Students had to respond on a five-point scale ranging from 1 (*totally incorrect*) to 5 (*totally correct*). The internal consistency was sufficient (Cronbach’s alpha = .67).

3.2.4 Covariates

Information on gender and school type was collected from school administrations. School type was used as a covariate because previous analyses of the *LISA 6* data have revealed that academic-track students show higher academic achievement as well as higher SES and cognitive ability than vocational-track students (see Leucht & Köller, 2016). Students’ SES was operationalized by the *International Socio-Economic Index of Occupational Status* (ISEI; Ganzeboom, De Graaf, Treiman, & De Leeuw, 1992). Data were collected from both parents, and the highest ISEI in the family (HISEI) was computed as an indicator of SES. Higher HISEI values indicate a high SES.

3.3 Statistical analyses

We ran a series of path analyses using M*plus* (Version 8.0; Muthén & Muthén, 1998–2017) to examine our research questions. The path analyses were conducted domain-specifically, that is, separately for the three domains and for GPA. Due to the different metrics of the dependent and independent variables, completely standardized solutions were inspected to investigate the predictive power of intelligence and motivation. We interpreted nonoverlapping 95 % confidence intervals as significant differences between the standardized regression coefficients. We decided to run manifest analyses without correcting for measurement errors because most measures had high reliabilities and some measures did not have any reliability information (grades and final exams).

All models were based on maximum likelihood estimation. Because of the hierarchical data structure with students nested in classes and schools, it was necessary to control for dependencies in the data. Thus, we took the data structure into account by computing robust estimates of the model parameters and their standard errors (Type = Complex in M*plus*; see Muthén & Satorra, 1995) in all models.

Missing data were a serious problem in the present investigation. Table 1 provides missing information for all the measures. Especially the student questionnaire, which was voluntary, showed substantial numbers of missings.[1] We thus decided to apply multiple imputation (MI) techniques to generate a large body of

1 Note that the voluntary participation of students in questionnaire studies is regulated by law, which often leads to substantial percentages of missing data in such studies. Participation in standardized tests is usually mandatory.

complete data sets. Since full data sets were available for the test measures (PVs), we ran nested imputations (e.g., Harel & Schafer, 2003; Rubin, 2003; Weirich, Haag, Hecht, Böhme, Siegle, & Lüdtke, 2014), that is, we generated 20 full data sets without missings for each of the five PV vectors, resulting in 100 data sets overall that formed the basis for all analyses presented in the Results section. We used M*plus* to analyze all data sets. M*plus*, however, only provides combined results for nonnested MI data sets. This can result in slightly biased standard errors of estimated parameters. Therefore, we decided to set a more conservative alpha level of $p < .01$.

Table 1: Descriptive statistics (means, standard deviations, and ranges for nonstandardized variables) and percentage of missing data in the present study

Variable	*M*	*SD*	Range	Percentage of missings
Mathematics				
Test	0.00	1.00		-
Grade	8.37	3.27	0 – 15	3.0
Final exam	6.47	3.55	0 – 15	18.6
Self-concept	2.43	0.88	1 – 4	45.1
Interest	2.14	0.88	1 – 4	45.2
English				
Test	0.00	1.00		-
Grade	8.78	2.74	0 – 15	3.2
Final exam	8.15	2.81	0 – 15	21.0
Self-concept	2.63	0.76	1 – 4	44.5
Interest	2.42	0.73	1 – 4	44.9
German				
Grade	8.69	2.51	0 – 15	3.0
Final exam	7.06	2.65	0 – 15	19.2
Self-concept	2.53	0.72	1 – 4	45.5
Interest	2.48	0.82	1 – 4	45.2
General				
Intelligence	0.00	1.00		-
Self-concept	3.17	0.65	1 – 4	45.0
Conscientiousness	3.65	0.68	1 – 5	55.6

4. Results

4.1 Descriptive and correlational findings

Table 1 provides descriptive information on all the variables, including the percentages of missing data. Correlational findings for the different domains and the general measures are presented in Tables 2 and 3. Large correlations were found between the different achievement measures, particularly between report card grades and final written exams. Intelligence had high correlations with standardized test results in English and the highest correlations with test results in mathematics. Domain-specific self-concept and interest had very strong associations, with correlations above .70. However, the correlations between self-concept and achievement variables were stronger than those between interest and achievement variables, thereby supporting findings from previous studies (e.g., Lotz et al., 2018; Marsh et al., 2005).

Concerning the general measures, intelligence, self-concept, and conscientiousness showed more or less the same relationships with GPA. The somewhat smaller coefficient between conscientiousness and GPA is due to the lower reliability of the conscientiousness scale.

Table 2: Correlations among variables; measures for English above the diagonal; for mathematics below the diagonal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Test (1)		.51	.56	-.07	.37	.18	.46	.47	.33
Grade (2)	.44		.77	.06	.15	.19	.21	.60	.40
Final exam (3)	.52	.70		.02	.22	.19	.27	.56	.34
Sex[1] (4)	-.45	.03	-.07		-.06	-.07	-.15	.05	.05
School type[2] (5)	.33	.13	.15	-.06		.24	.37	.07	.03
HISEI (6)	.14	.10	.13	-.07	.24		.11	.13	.11
Intelligence (7)	.64	.31	.37	-.15	.37	.11		.06	.01
Self-concept (8)	.44	.59	.57	-.15	.05	.01	.29		.71
Interest (9)	.36	.44	.45	-.16	-.01	.00	.23	.83	

Note. [1]Reference is female; [2]Reference is academic-track school; analyses were based on 100 nested imputed data sets; combined coefficients; coefficients above .09 and below -.09 are significant ($p < .01$).

Table 3: Correlations among variables; measures for German above the diagonal; general measures below the diagonal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Grade (1)		.64		.10	.16	.15	.19	.43	.31
Final exam (2)				.09	.16	.12	.21	.38	.21
GPA (3)									
Sex[1] (4)			-.05		-.06	-.07	-.15	.25	.19
School Type[2] (5)			-.14			.24	.37	.04	-.04
HISEI (6)			-.18				.11	.03	.00
Intelligence (7)			-.30					-.06	-.11
Self-concept (8)			-.30	-.19	.03	-.12	.20		.78
Interest (9)									
Conscientiousness (10)			-.24	.24	-.07	-.01	-.06	.18	

Note. [1]Reference is female; [2]Reference is academic-track school; analyses were based on 100 nested imputed data sets; combined coefficients; coefficients above .09 and below -.09 are significant ($p < .01$).

4.2 Findings of path analyses

Table 4 provides the results of the path analyses that were conducted separately for each domain and for the general measures. The findings represent the combined results from 100 imputed data sets (see Method section).

Table 4: Intelligence and motivation as predictors of achievement (score-card grade, final written exam, and standardized test); standardized regression coefficients/standard errors from path analyses

	Dependent Variable								
	English			Mathematics			German		GPA
	Grade	Exam	Test	Grade	Exam	Test	Grade	Exam	
Intelligence	.17*/.02	.20*/.02	.35*/.02	.14*/.02	.20*/.02	.48*/.02	.19*/.02	.21*/.02	-.26*/.02
Self-concept	.61*/.02	.60*/.02	.40*/.02	.66*/.03	.55*/.03	.24*/.03	.45*/.03	.52*/.03	-.21*/.02
Interest	-.05/.03	-.10*/.03	.04/.03	-.11*/.03	-.04/.03	.00/.03	-.03/.03	-.18*/.03	
Conscientiousness									-.20*/.02
Covariates									
Sex[1]	.06*/.02	.04*/.02	-.03/.02	.13*/.02	.04/.02	-.34*/.01	.04/.02	.04*/.02	-.09*/.02
School-type[2]	.03/.02	.09*/.02	.21*/.02	.03/.03	.03/.03	.12*/.02	.05/.03	.04/.03	-.02/.03
HISEI	.10*/.02	.08*/.03	.03/.02	.08*/.02	.10*/.02	.04/.02	.11*/.02	.08*/.02	-.13*/.02
R^2	51.1	39.0	44.4	39.4	38.3	61.2	25.0	22.4	22.3

Note. * $p < .01$; R^2: Percentage of explained variance; [1]Reference is female; [2]Reference is academic-track school; analyses were based on 100 nested imputed data sets; combined coefficients.

English. Both intelligence and motivational measures had substantial effects on all three measures. However, once we controlled for English self-concept, the effects of interest became close to zero or even negative, which can be explained by multicollinearity (high correlation between self-concept and interest). Intelligence had the largest effect on test scores and effects on grades and written exams were smaller. The picture for self-concept is quite different: it had very large effects on grades and written exams but somewhat smaller effects on test scores. Although the predictive power of intelligence and self-concept was nearly the same for test scores, self-concept clearly outperformed intelligence in the case of grades and written exams. The effects of all the covariates were small: Female students slightly outperformed male students; high SES-students performed a little better than low SES-students and students from academic-track schools outperformed those from vocational-track schools.

Mathematics. To some extent the picture in mathematics corresponds to the one in English with one exception: intelligence outperformed self-concept in predicting standardized tests scores. The substantial effect of gender on test scores indicates that male students outperformed female students. Interestingly, the gender effect for grades indicates that female students (after controlling for all other measures in Table 4) got better grades than male students.

German. Again, self-concept outperformed intelligence but both predictors had substantial effects on grades and written exams. The effects for self-concept were a little bit smaller than in English and mathematics. The negative effect of interest on written exams is again a result of multicollinearity. Regarding the covariates, a small but significant effect was found in favor of high SES-students. Overall, less variance was explained by all of the predictors.

GPA[2]. Intelligence, self-concept, and conscientiousness had effects of similar size. Again, the effects of the covariates were small, with female students having slightly better GPAs than male students and high SES-students getting better GPAs than low SES-students. Congruent to German, the amount of explained variance was smaller than in English and mathematics.

In summary, we found evidence that, in most cases, motivational variables outperformed intelligence in predicting achievement. However, intelligence had additional substantial effects on all achievement measures; however, this may have been due to the fact that the data were collected in a cognitively selected sample with reduced variance in intelligence.

5. Discussion

The major goal of the present study was to expand our knowledge of the differential predictive value of intelligence and motivation for academic achievement in upper secondary schools. We considered domain-specific academic outcomes in

2 Note that high values for GPA indicate low overall achievement (1 = *excellent*; 4 = *poor*).

English, mathematics, and German, as well as GPA to find out whether the pattern of results was moderated by the domain, the type of achievement indicator, and the domain-specificity. Concerning the different domains, our results indicate relatively high consistency in the pattern of findings: Domain-specific self-concept was the dominant predictor for written exams and grades. These regression weights became smaller when looking at standardized tests. Interest, as the second domain-specific motivational predictor, was positively correlated with all achievement measures. However, in the path analyses, the regressions weights came close to zero or were even slightly negative. Intelligence showed consistently moderate effects on grades and written exams. In line with previous literature (Lotz et al., 2018), the effects were higher with respect to standardized tests in English and in mathematics. Regarding the prediction of GPA, academic self-concept, conscientiousness, and intelligence showed effects of similar size. Thereby, it is important to note that a personality trait was predictive beyond the effect of academic self-concept. Even though our study was limited to a cross-sectional design and thus did not make it possible to study effects on changes in achievement, it provides some new insights and raises questions about the detailed relationships of the different predictors with academic achievement.

5.1 Predictive power of intelligence

Many previous studies have consistently shown the predictive power of intelligence for academic achievement. Our study replicates these findings. However, the effects were relatively small for grades and written exams. One reason for these lower coefficients could be the selectivity of our sample. As we only sampled students at the end of upper secondary school, more than 50 % of the age cohort was not part of the sample because they left secondary school after Grade 9 or Grade 10 to commence vocational education. Research findings suggest that students who leave school earlier have lower intelligence scores than those who attend upper secondary schools (see the recent study by Guill, Lüdtke, & Köller, 2017, on intelligence differences in different school types in Germany). Therefore, a shrinkage of variance in intelligence may have been the reason for the lower correlations and regression weights.

The effects of intelligence on standardized tests scores were somewhat higher and support the argument that the type of achievement measure moderates the predictive effect of intelligence. Although this is fully in line with previous work (Lotz et al., 2018), we would like to point out, as mentioned above, that some authors (e.g., Kaufman et al., 2012; Saß et al., 2017) have recently argued that academic achievement tests partly measure intelligence. If this is true, one might argue that the higher predictive power of intelligence on standardized tests is merely a consequence of the fact that both tests, at least to some extent, measure the same construct.

5.2 Predictive power of domain-specific self-concept and interest

A recent study by Lotz et al. (2018) revealed that mathematics-specific self-concept had substantial effects on mathematics achievement, but that mathematics-specific interest did not have any effects once self-concept was controlled for. Our results strongly support this finding and confirm the argument by Lotz et al. (2018) that, in the EVT, self-concept is the dominant predictor when achievement is the dependent variable, and that interest seems to be more important when academic choices are predicted. Another explanation for the nonsignificant to negative effects of interest might be that we only used cross-sectional data with performance measures. Domain-specific interest with its intrinsic character might be more important when learning (i.e., change in achievement) is the central outcome in longitudinal studies. Köller et al. (2001) provided support for the assumption that mathematics interest predicts mathematics learning when highly interested students have the chance to opt for advanced courses in which they have more mathematics lessons of higher instructional quality while low-interest students opt for basic courses with a low quantity and quality of instruction.

Compared to many previous studies, we found even higher standardized regression coefficients for domain-specific self-concept. This was particularly true for grades and final exams. Again, we have to mention that our sample was cognitively positively selected, that is, our sample only comprised students from upper secondary schools. As a consequence of this cognitively more homogeneous group, motivation in terms of domain-specific self-concept became a more important predictor of students' academic outcomes. Obviously, high self-concept students put more effort into their final written exams. The strong correlations between grades and self-concept might not only be the result of higher performance but also of more participation and supportive behavior during classroom lessons. There is much evidence (e.g., Kaiser, Retelsdorf, Südkamp, & Möller, 2013) that teachers reinforce such behavior by providing better grades. We are, however, aware that the coefficients of self-concept typically become smaller in longitudinal research when prior grades are included as predictors of subsequent grades (see Limitations section). Finally, our results provide support for our hypothesis that the type of achievement measure moderates the predictive power of motivation, with the finding that the regression weights of self-concept were somewhat lower for standardized tests.

5.3 Prediction of GPA by general measures

Finally, we analyzed the effects of academic self-concept, conscientiousness, and intelligence on GPA. Because GPA represents a global measure of academic achievement, we did not focus on domain-specific motivational factors as predictors. Using self-concept and conscientiousness as global motivational measures, we found evidence for their predictive power above and beyond intelligence. It is noteworthy, however, that the motivational measures showed effect sizes similar to those for intelligence. In contrast to the domain-specific achievement measures, motivation was not a clearly predominant predictor. These findings suggest that the predominance of motivation disappears once intelligence and motivation are measured at the same aggregation level (aggregated across all domains).

5.4 Limitations

Although our study provides new insights into the relationship between intelligence, motivation, and achievement, there are some limitations that require additional research. First of all, the cross-sectional design did not allow an analysis of the effects on changes in achievement. Much evidence from previous research shows that once prior achievement is controlled for, the effects of all motivational and other cognitive predictors decrease substantially but still remain positive and significant (e.g., Marsh et al., 2005). To inspect such effects in our study, we ran some additional analyses predicting achievement in final written exams in mathematics, German, and English while controlling for report card grades. The effects of domain-specific self-concept and intelligence decreased but remained significant ($p < .01$; standardized path coefficients: mathematics/German/English: self-concept: .199/.279/.210; intelligence: .128/.104/.096), again providing evidence for their predictive power.

A methodological limitation of the present study is that all analyses were based on multiple data sets from nested imputation procedures but that the software package used (M*plus*) ignores the nested character of all data sets. This could have resulted in somewhat biased estimates of standard errors. Being aware of this problem, we used a more conservative alpha level ($p < .01$). Our impression, however, is that much more research in handling data from the multiple nested imputation of missing values is necessary. We nevertheless feel that our strategy is defendable due to the large amount of missing data from the student questionnaire and the fact that plausible values were available for the standardized tests.

Finally, we restricted our analyses to manifest variables and did not control for measurement error. We are sure that the pattern of results would have been stable even after controlling for measurement error. As the reliabilities of most measures were very high, one would expect slightly higher regression coefficients after controlling for measurement error but not a substantial change in the overall picture of findings.

5.5 Conclusions

A long research tradition has addressed the effects of intelligence and motivation on academic achievement. However, there is still a lack of studies that simultaneously analyze different domains, different achievement outcomes, and different motivation measures. The current study thus contributes substantially to the literature. The strong effects of domain-specific self-concept on all the different achievement measures highlight the great importance of motivation in explaining differences in achievement at the end of upper secondary school. Although the effects of intelligence were somewhat smaller, our findings also underline the important role that intelligence plays in the school context beyond motivation. This is true, however, for different domains and different achievement measures.

References

Adams, R. J, Wu, M. L., & Wilson, M. R. (2012). *ACER ConQuest 3.0.1*. Camberwell, England: ACER.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*, 359-372.

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition. Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*, 165–176.

Benson, N. F., Kranzler, J. H., & Floyd, R.G. (2016). Examining the integrity of measurement of cognitive abilities in the prediction of achievement: Comparisons and contrasts across variables from higher-order and bifactor models. *Journal of School Psychology, 58*, 1–19.

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, *17*, 69–81.

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*, 425–453.

Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and social Psychology, 11,* 492–511.

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13–21.

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality, 10*, 303–336.

Dumfart, B., & Neubauer, A. C. (2016). Conscientiousness is the most powerful noncognitive predictor of school achievement in adolescents. *Journal of Individual Differences*, *37*, 8–15.

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: Freeman.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501–519.

Fink, M. B. (1962). Objectification of data used in underachievement self-concept studies. *California Journal of Educational Research, 13*, 105–112.

Fleckenstein, J., Leucht, M., Pant, H. A., & Köller, O. (2016). Proficient beyond borders: Assessing non-native speakers in a native speakers' framework. *Large-scale Assessments in Education, 4*(19).

Ganzeboom, H. B. G., Graaf, P. M. de, & Treiman, D. J., & De Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, *21*, 1–56.
Geisler-Brenstein, E., Schmeck, R. R., & Hetherington, J. (1996). An individual difference perspective on student diversity. *Higher education, 31*(1), 73–96.
Gough, H. G., & Fink, M. B. (1964). Scholastic achievement among students of average ability, as predicted from the California Psychological Inventory. *Psychology in the Schools, 1*(4), 375–380.
Guay, F., Ratelle, C. F., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences, 20*, 644–653.
Guill, K., Lüdtke, O. & Köller, O. (2017). Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study. *Learning and Instruction, 47*, 43–52.
Harel, O., & Schafer, J. L. (2003). *Multiple imputation in two stages*. Washington DC: Paper presented at the Proceedings of Federal Committee on Statistical Methodology Research Conference.
Heller, K. A., & Perleth, C. (2000). *KFT 4-12+ R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [Cognitive ability test for Grade 4]. Göttingen, Germany: Beltz.
Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen* [Self-confidence and academic achievement]. Göttingen, Germany: Hogrefe.
Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality*, *52*, 29–36.
Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology, 46*, 116–127.
Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
Jopt, U. J. (1978). *Selbstkonzept und Ursachenerklärung in der Schule* [Self-concept and causal attribution in school]. Bochum, Germany: Kamp.
Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction, 28*, 73–84.
Kampa, N., Köller, O., Schmidt, F. T. C., & Leucht, M. (2016). Kompetenzen im Fach Mathematik. [Competencies in mathematics]. In M. Leucht, N. Kampa, & O. Köller (Eds.), *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* (pp. 119–143). Münster, Germany: Waxmann.
Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence, 40*, 123–138.
Kling, K. C., Noftle, E. E., & Robins, R. W. (2013). Why do standardized tests underpredict women's academic performance? The role of conscientiousness. *Social Psychological and Personality Science*, *4*, 600–606.
KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany] (2002a). Einheitliche Prüfungsanforderungen in der Abiturprüfung Englisch [Standardized requirements for final examinations in English at the end of upper secondary school] (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 24.05.2002). Bonn, Germany: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Retrieved from http://www.kmk.org/fileadmin/Dateien/ veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Englisch.pdf
KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [Standing Conference of the Ministers of Education and Cultural

Affairs of the Länder in the Federal Republic of Germany] (2002). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Mathematik* [Standardized requirements for final examinations in mathematics at the end of upper secondary school] (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 24.05.2002). Bonn, Germany: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Retrieved from http://www.kmk.org/fileadmin/ Dateien/ veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Mathe.pdf

Köller, O., Knigge, M., & Tesch, B. (Eds.). (2010). *Sprachliche Kompetenzen im Ländervergleich* [Verbal competencies in a national comparison among federal states]. Münster, Germany: Waxmann.

Köller, O., Schnabel, K., & Baumert, J. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education, 32*, 448–470.

Köller, O., & Trautwein, U. (2004). Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien [English achievement of students from academic track and vocational upper secondary schools]. In O. Köller, R. Watermann, U. Trautwein, & O. Lüdtke (Eds.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (pp. 285–326). Opladen, Germany: Leske + Budrich.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, *51*(4), 472–477.

Kriegbaum, K., Jansen, M., & Spinath, B. (2015). Motivation: A predictor of PISA's mathematical competence beyond intelligence and prior achievement. *Learning and Individual Differences, 43*, 140–148.

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences, 42*, 441–451.

Leucht, M., Fleckenstein, J., & Köller, O. (2016). Erreichen kriterialer Leistungsstandards in der ersten Fremdsprache Englisch. [Reaching proficiency levels in English as a foreign language]. In M. Leucht, N. Kampa, & O. Köller (Eds.), *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* (pp. 171–199). Münster, Germany: Waxmann.

Leucht, M., Kampa, N., & Köller, O. (Eds.) (2016). *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* [Academic achievement at the end of upper secondary school. A comparison of vocational and academic upper secondary schools]. Münster, Germany: Waxmann.

Leucht, M., & Köller, O. (2016). Anlage und Durchführung der Studie [Design and method of the study]. In M. Leucht, N. Kampa, & O. Köller (Eds.), *Fachleistungen beim Abitur. Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein* (pp. 79 – 98). Münster, Germany: Waxmann.

Lotz, C., Schneider, R., & Sparfeldt, J. (2018). Differential relevance of intelligence and motivation for grades and competence tests in mathematics. *Learning and Individual Differences, 65*, 30–40.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*, 133–163.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397–416.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality, 74*, 403–456.

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research, 79*, 1129–1176.

Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology, 84*, 591–611.

Murray, H. A. (1938). *Explorations in personality*. Oxford, England: Oxford University Press.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25*, 267–316.

Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E., & Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for Educational Research Online, 5*, 80–109.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*, 322–338.

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K) [Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality]. *Diagnostica*, *51*, 195–206.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118–137.

Rubin, D.B. (2003). Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica, 57*, 3–18.

Saß, S. Kampa, N., & Köller, O. (2017). The interplay of g and mathematical abilities in large-scale assessments across grades. *Intelligence, 63*, 33–44.

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Hillsdale, NJ: Erlbaum.

Schmidt, F. T. C., Fleckenstein, J., Retelsdorf, J., Eskreis-Winkler, L., & Möller, J. (2017). Measuring grit: A German validation and a domain-specific approach to grit. *European Journal of Psychological Assessment*, Advanced Online Publication, 1–12. Retrieved from https://econtent.hogrefe.com/doi/abs/10.1027/1015-5759/a000407

Schneider, F. W., & Green, J. E. (1977). Need for affiliation and sex as moderators of the relationship between need for achievement and academic performance. *Journal of School Psychology, 15*, 269–277.

Schneider, W., & Niklas, F. (2017). Intelligence and verbal short-term-/working memory: Their interrelationships from childhood to young adulthood and their impact on academic achievement. *Journal of Intelligence, 5*, 26.

Sorić, I., Penezić, Z., & Burić, I. (2017). The Big Five personality traits, goal orientations, and academic achievement. *Learning and Individual Differences*, *54*, 126–134.

Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence, 34*, 363–374.

Spinath, B., Freudenthaler, H. H., & Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Personality and Individual Differences, 48*, 481–486.

Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Eds.) (2016). *IQB Bildungstrend 2015: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. [IQB Educational Trend 2015: Second national assessment on Language competencies at the end of grade 9]. Münster, Germany: Waxmann.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19*, 80–90.

Stumm, S., von Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*, 574–588.

Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology, 97*, 1115–1128.

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*, 438–456.

Weirich, S., Haag, N., Hecht, M., Böhme, K., Siegle, T., & Lüdtke, O. (2014). Nested multiple imputation in large-scale assessments. *Large-scale Assessments in Education, 2*(9).

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