

## Are more conscientious seventh- and ninth-graders less likely to be retained? Effects of Big Five personality traits on grade retention in two different age cohorts<sup>☆, ☆☆, ☆☆☆</sup>

Andrea Westphal<sup>\*</sup>, Miriam Vock, Rebecca Lazarides

University of Potsdam, Germany

### ARTICLE INFO

#### Keywords:

Grade retention  
Personality  
Gender  
Socioeconomic status  
Educational large-scale study  
Educational achievement

### ABSTRACT

Previous research has identified students' personality traits, especially conscientiousness, as highly relevant predictors of academic success. Less is known about the role of Big Five personality traits in students when it comes to teachers' decisions about students' educational trajectories and whether personality traits differentially affect these decisions by teachers in different grade levels. This study examines to what extent students' Big Five personality traits affect teacher decisions on grade retention, looking at two cohorts of 12,146 ninth-grade and 6002 seventh-grade students from the German National Educational Panel Study. In both grade levels, multilevel logistic mediation models show that students' conscientiousness indirectly predicts grade retention through the assignment of grades by teachers. In the ninth-grade sample, students' conscientiousness was additionally a direct predictor of retention, distinct from teacher-assigned grades. We discuss potential underlying mechanisms and explore whether teachers base their decisions on different indicators when retaining seventh-grade students or ninth-grade students.

### Introduction

In many countries in North America and Western Europe, > 10% of all 15-year-old students have been retained at least once in their school career (e.g., Ikeda & García, 2014). In the case of the United States, where around 2.3% of all students are retained every year and 14% of 15-year-old students have previously been retained, there are estimates suggesting that the cost of retaining so many students is higher than \$12 billion per year (West, 2012). Whether or not grade retention is beneficial to the academic and socioemotional development of students is a controversial and widely discussed issue. While there is some indication that students may benefit from grade retention in terms of teacher-assigned grades and student self-efficacy (Marsh et al., 2017; Wu, West, & Hughes, 2010), other studies show no positive or even detrimental effects of grade retention on achievement, academic self-

concept and school-belonging (Im, Hughes, Kwok, Puckett, & Cerda, 2013; Martin, 2011).

The decision to have a student repeat a grade level or not is usually made by teachers (Tanner & Galis, 1997) who are expected to take the student's course grades into account in making this determination. In Germany, where data for this study were collected, the laws of the federal states regulate the details for repeating a class level (e.g., School Law of the Federal State North Rhine-Westphalia, 2005). While empirical studies have confirmed that retained students had much lower teacher-assigned grades than their classmates (Ehmke, Drechsel, & Carstensen, 2008), there is also some evidence suggesting that boys and students with lower socioeconomic status (SES) have a substantially higher risk of being retained than both girls and students with higher SES (Entwisle, Alexander, & Olson, 2007; Klapproth & Schaltz, 2015). In addition, surveys have shown that teachers consider not only their

<sup>☆</sup> We thank Ben Fergusson for his editorial assistance.

<sup>☆☆</sup> This study uses data from the National Educational Panel Study (NEPS): Starting Cohort Grade 5 (doi:<https://doi.org/10.5157/NEPS:SC3:6.0.1>) and Starting Cohort Grade 9 (doi:<https://doi.org/10.5157/NEPS:SC4:7.0.0>). From 2008 to 2013, NEPS data were collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

<sup>☆☆☆</sup> Our research was supported by the Deutsche Forschungsgemeinschaft (DFG, grant number WE 6191/1-1). The DFG had no involvement in the collection, analysis and interpretation of data or in the writing of this manuscript.

<sup>\*</sup> Corresponding author at: Department of Education, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.

E-mail address: [andrea.westphal@uni-potsdam.de](mailto:andrea.westphal@uni-potsdam.de) (A. Westphal).

students' achievement in class, but also other non-achievement factors, such as their students' effort and motivation in class, as criteria on which they base their decision to retain or promote a student (Bless, Schüpbach, & Bonvin, 2004; Tomchin & Impara, 1992; Witmer, Hoffman, & Nottis, 2004). Students' Big Five personality traits—i.e., automatic thoughts, feelings and behaviors that are relatively stable across time (Damian, Spengler, Sutu, & Roberts, 2018) and have previously been linked to students' academic success (e.g., Spengler, Lüdtke, Martin, & Brunner, 2013)—may also be associated with a student's risk of being retained. Students with less favorable personality traits, for instance less conscientious students, may receive lower grades and may therefore be at higher risk for being retained. Conversely, teachers may take students' personality favorably into account when deciding to retain a student. A less conscientious student may be more likely to be retained than a more conscientious student with similar teacher-assigned grades.

Thus, there is some indication that teachers take factors other than achievement into account when retaining students. Nonetheless, the role of students' Big Five personality traits in teachers' decisions to retain a student has not yet been investigated. Based on a large longitudinal multi-cohort sample from Germany, the present study examines whether grade retention can be predicted by students' personality above and beyond their teacher-assigned grades, gender, and SES. The study focused on retention in Grade 7 and Grade 9, and thus on grade levels in which students in Germany frequently are retained (German Federal Statistical Office, 2017).

#### *Frequency of grade retention*

Within the Organization for Economic Co-operation and Development (OECD), some countries, such as the United Kingdom, make almost no use of retention (2% of 15-year-old students have previously been retained), while other countries—such as France, Portugal, and Spain—have higher retention rates (35% to 37% of 15-year-old students retained), especially in grade levels that precede critical examinations (Ikeda & García, 2014). Retention rates in the United States and Germany tend to lie in between (Ikeda & García, 2014). In Germany, the frequency of grade retention has slightly decreased over the last decade. In the school year 2015/16, 2.3% of German school students repeated a grade (German Federal Statistical Office, 2017), while the rate was 2.9% in the school year 2003/04 (German Federal Statistical Office, 2004). The highest rates of grade retention in Germany have been found for Grades 7 to 12 (German Federal Statistical Office, 2017). One potential reason for the higher rates of retention in these grade levels may be that they mark a vulnerable developmental period in students' lives as they transition from childhood to adolescence, a period characterized by age-related changes and challenges (Masten, Obradović, & Burt, 2006). In adolescence, students begin to detach from their parents and develop close relationships with peers (Bandura, 1964; Erikson, 1959; Havighurst, 1948). Related to these changes, students also experience difficulties in focusing on learning and report lower levels of academic effort and persistence (Pajares & Graham, 1999). Consequently, one of the major difficulties for teachers during this period is adolescent students' disengagement from schooling (Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). These disengaging behaviors include behavioral disengagement (e.g., not attending class, not completing schoolwork, lower levels of effort and participation), emotional disengagement, and cognitive disengagement (e.g., less self-regulated learning) (Wang & Holcombe, 2010). These kinds of disengaging behaviors negatively affect teacher-assigned grades (Wang & Holcombe, 2010) and may also explain why students in these grade levels are particularly vulnerable to being held back a year. This means that research into grade retention seems especially crucial when it comes to these grade levels.

#### *Effects of grade retention on students' academic development*

Debates about the usefulness of grade retention are underpinned by contrasting theories and beliefs about child development. These theories differ in their understanding of students' individual needs and in the criteria applied to decide whether or not a student should learn with their same age peers (for an overview on these theories, see Martin, 2009; Meisels, 1998). Proponents of grade retention would argue, from a nativist perspective, that each student develops at his or her own pace, which ideally should be similar to that of the student's learning environment and peers—otherwise a student is at higher risk of academic failure (Martin, 2009; Meisels, 1998). Following this rationale, grade retention may be a good way to support students whose development proceeds more slowly than that of their peers (Martin, 2009). In contrast, opponents of grade retention may adhere to the environmental perspective, which posits that children develop by engaging with their environment—which is therefore considered responsible for speeding up a child's development (Meisels, 1998). This perspective would suggest promoting a student and instead altering the environment in order to meet the student's needs (Martin, 2009).

Empirical research on the benefits and drawbacks of grade retention for student development does not strongly oppose grade retention, but does not encourage the practice either. Based on his meta-analysis, Jimerson (2001) concluded that grade retention has substantial negative effects on students' achievement in reading, language and mathematics, as well as on their emotional adjustment and school attendance. In a more recent meta-analysis, Allen, Chen, Willson, and Hughes (2009) showed that negative effects of grade retention on student achievement are more pronounced in investigations applying a less advanced research design (effect size of  $-0.30$ ), while studies using more adequate comparison groups and statistical control variables found no statistically significant or practically meaningful effect of retention on student achievement (effect size of  $0.04$ ). The authors also showed that an academic advantage that may occur immediately after grade retention is followed by a “downward trajectory” (p. 493) in the ensuing years. Similar effects of positive results fading away have been found for students' engagement and school-belonging (Im et al., 2013). There is some recent indication that retained students may benefit in terms of a higher academic self-concept and less anxiety in class (Marsh, 2016; Marsh et al., 2017). It has, however, been pointed out that the time of measurement of the baseline may be crucial because socioemotional characteristics may be at rock bottom immediately before retention, and benefits of grade retention are therefore likely overestimated when the wrong baseline is chosen (Kretschmann, Vock, Lüdtke, Jansen, & Gronostaj, 2017). Given the substantial financial costs and the emotional burden for retained students, the largely non-significant differences between retained and promoted students “[call] into question the educational benefits of grade retention policies” (Allen et al., 2009, p. 493).

#### *The role of Big Five personality traits in students in their success in school*

The *Big Five personality traits* in adolescents, i.e., openness, conscientiousness, extraversion, agreeableness, and emotional stability, seem to be highly relevant in their (dis)engagement with school and consequently their educational success. Personality traits are a crucial factor, for instance, in adolescents' attendance (Lounsbury, Steel, Loveland, & Gibson, 2004) and academic motivation (McGeown et al., 2014), as well as in their scores on standardized achievement tests (Spengler et al., 2013) and the grades they receive in school (Laidra, Pullmann, & Allik, 2007; Spengler et al., 2013). When it comes to teacher-assigned grades—which are highly relevant for grade retention—different studies have identified different personality traits as being relevant (openness: e.g., Steinmayr & Spinath, 2007; conscientiousness: e.g., Laidra et al., 2007, Spengler et al., 2013, Spinath, Freudenthaler, & Neubauer, 2010; extraversion: e.g., Spinath et al.,

2010, Steinmayr & Spinath, 2007; agreeableness: e.g., Steinmayr & Spinath, 2007; emotional stability: e.g., Spinath et al., 2010). Yet, there is some indication that students' conscientiousness seems to be of particular relevance for teacher-assigned grades in secondary school (Laidra et al., 2007; Spengler et al., 2013). Students that are more conscientious exhibit more self-control (MacCann, Duckworth, & Roberts, 2009). Therefore, it would be plausible to assume that they might spend more time learning (Poropat, 2014) and might in turn receive better grades. In the light of this, we would expect that the Big Five personality traits in secondary-school students—especially conscientiousness—would be associated with grade retention. It is therefore surprising that the role of Big Five personality traits in the grade retention of adolescents has not yet been studied.

### *The mechanisms of grade retention*

In Germany, the decision on whether a student should be retained is based on the teacher-assigned grades a student has received over the course of a school year (e.g., *School Law of the Federal State North Rhine-Westphalia*, 2005). The details vary slightly across the sixteen federal states and across the different school tracks. However, in all cases, grades in mathematics and German are crucial in this decision (e.g., *School Law of the Federal State Baden-Wuerttemberg*, 2004). In this study, we focus, therefore, on students' grades in these two subjects when examining the effects of personality traits and achievement on grade retention.

It may be justified to include aspects other than teacher-assigned grades in decisions on retention, however, if we take education legislation into account at the state level. For instance, the federal states allow a meeting of all of the relevant student's teachers to promote students, even if that student does not achieve the teacher-assigned grades required to be promoted (e.g., *School Law of the Federal State Baden-Wuerttemberg*, 2004). This is possible if the meeting of all of the relevant student's teachers declares, with a two-thirds majority, that the student will likely succeed in the higher grade-level (e.g., *School Law of the Federal State Baden-Wuerttemberg*, 2004).

Accordingly, research examining the differences between retained and promoted students has shown that students who are retained exhibit less favorable teacher-assigned grades in mathematics and in the main linguistic subject taught in the school (e.g. German in Germany etc.) (Dauber, Alexander, & Entwisle, 1993; Ehmke et al., 2008). But other factors—such as students' gender, SES, academic intentions, school absence, and motivational factors (e.g., homework completion, class participation, task persistence, and school enjoyment)—are also associated with grade retention (Dauber et al., 1993; Ehmke et al., 2008; Entwisle et al., 2007; Huang, 2014; Martin, 2009; McCoy & Reynolds, 1999). Some of these characteristics play a role in grade retention beyond teacher-assigned grades. For instance, boys are more likely to be retained than girls, even when their teacher-assigned grades are similar (Dauber et al., 1993; McCoy & Reynolds, 1999).

In addition to this, teachers report that they also take characteristics other than achievement into account, for instance the student's effort, motivation and classroom behavior (Bless et al., 2004; Tomchin & Impara, 1992; Witmer et al., 2004), when they decide on a student's retention. In the qualitative study of Tomchin and Impara (1992), teachers explained why they took such factors into account in their retention decisions. Some teachers thought that students who did not show enough effort in class, were lazy, unmotivated or poorly organized should be punished, and viewed retention as a means of accomplishing this. Other teachers, in contrast, reported that they viewed non-cognitive factors as predictive of whether a low-achieving student would succeed in the next grade level (Tomchin & Impara, 1992). As such, teachers' views on the Big Five personality traits in adolescents might also have a positive effect when deciding whether to retain a student.

### *The process by which students' Big Five personality traits affect grade retention*

Because retention decisions, at least in Germany, should primarily be based on teacher-assigned grades, the question arises as to whether students' Big Five personality traits affect teachers' retention decision by impacting teacher-assigned grades or, else, influencing teachers' retention decisions over and beyond teacher-assigned grades. Thus, students' Big Five personality traits may affect grade retention both directly and indirectly via teacher-assigned grades. Because of their close relations to grades, we expect that students' Big Five personality traits, especially conscientiousness, would be indirectly linked to grade retention via such teacher-assigned assessments. In other words, a less conscientious student may receive lower grades (e.g., Laidra et al., 2007; Spengler et al., 2013; Spinath et al., 2010) and may therefore have a greater risk of being retained.

Furthermore, teachers may favorably take students' personality traits into account when deciding whether a certain student should be retained or promoted. In Germany, state laws and regulations give teachers similar latitude. Student personality could therefore be directly linked to grade retention (after controlling for differences in teacher-assigned grades). Given that teachers reported that they considered students' work habits or laziness in their retention decisions (Tomchin & Impara, 1992; Witmer et al., 2004), we broadly expect to find that adolescents' conscientiousness has a direct effect on grade retention. Thus, teachers may promote a more conscientious student while retaining a student with similar grades, who, however, is less conscientious. Despite this, we will still look for direct effects of all Big Five personality traits in our analyses.

### *The present study*

We wanted to shed light on the question of whether or not adolescents who exhibit specific Big Five personality traits are at higher risk of being retained and, if so, whether this higher risk could solely be attributed to less favorable teacher-assigned grades. Using a large sample of German seventh-grade and ninth-grade students—surveyed at the beginning of the school year and tracking both their retention and promotion status at the end of the school year—we examined to what extent students' Big Five personality traits predicted grade retention over and beyond teacher-assigned grades. We focused on the following research questions:

1. To what extent do students' Big Five personality traits relate to grade retention?

Given previous findings on the relation between students' personality and teacher-assigned grades, as well as the fact that retention decisions are supposed to be primarily based on teacher-assigned grades, we explored the effects of personality traits on grade retention and expected to find substantial effects of students' conscientiousness on grade retention (Hypothesis 1).

2. To what extent do teacher-assigned grades relate to grade retention?

We hypothesized that grades in both mathematics and German would closely relate to grade retention (Hypothesis 2).

3. To what extent are the effects of students' Big Five personality traits on grade retention mediated by teacher-assigned grades?

Again, we expected to find substantial direct and indirect effects of students' conscientiousness on grade retention. Thus, we posited that less conscientious students receive lower grades and are therefore more likely to be retained (Hypothesis 3a). In addition, we hypothesized that less conscientious students are more likely to be retained than more

conscientious students with similar teacher-assigned grades, because teachers take students' conscientiousness favorably into consideration when deciding whether or not a student will succeed in the next grade level (Hypothesis 3b).

In order to differentiate direct and indirect effects of students' Big Five personality traits on grade retention via teacher-assigned grades, we specified a mediation model. We included students' gender, standardized test scores in mathematics and German, SES, and school absence as control variables in order to determine the incremental validity of students' Big Five personality traits in predicting grade retention. To simultaneously control the effects of these student characteristics on teacher-assigned grades, we also included indirect paths of these control variables on grade retention via teacher-assigned grades. We conducted a multilevel analysis to model the hierarchical data structure with students being nested in classes; this allowed us to additionally control for the average achievement in class and the academic track, both of which had previously been associated with teacher judgments and grade retention (see German Federal Statistical Office, 2017; Westphal et al., 2016).

## Method

### Samples

We used data from the National Educational Panel Study (NEPS; Blossfeld, Roßbach, & von Maurice, 2011), a longitudinal multi-cohort study examining educational processes and competence development. To this end, data from two samples of secondary school students from all sixteen German federal states is available.

#### Sample 1

Sample 1 comprised  $N = 16,425$  ninth-grade students in general and remedial secondary schools who were first surveyed when they were in the ninth grade. For our analyses, we selected students from general secondary schools ( $n = 15,239$ ) for whom data on their retention and promotion status, respectively, at the end of ninth-grade were available ( $n = 12,255$ ). We further excluded students with outlier values in days absent from school<sup>1</sup> ( $z$ -standardized values  $> 3.29$ ; 27 or more days absent from school). Our final sample therefore consisted of  $N = 12,146$  students. Of these students, 397 students were retained in Grade 9. We used data from the first, second and third panel wave, which took place at the beginning (November 2010 to January 2011; Wave 1) and end of ninth grade (May 2011 to July 2011; Wave 2) and in 10th grade (March 2012 to May 2012; Wave 3). In Wave 1 (January 2011), students were on average 14.7 years old ( $SD = 0.67$ ), 51.2% were female.

#### Sample 2

Sample 2 included  $N = 8317$  seventh-grade students in general and remedial secondary schools who had first been surveyed when they were in fifth grade. We selected students from general secondary schools ( $n = 7394$ ) who were surveyed along with their classmates<sup>2</sup> and for whom data on their retention or promotion status at the end of seventh-grade were available ( $n = 6055$ ). We further excluded students with outlier values in days absent from school ( $z$ -standardized values  $> 3.29$ ; 19 or more days absent from school). Thus, our final sample consisted of  $N = 6002$  students. Of these students, 131 students were retained in grade 7. We relied on data from the third and fourth

<sup>1</sup> In some cases, the reason for retaining students was their long absence from school (e.g., because of a severe illness), in which case individual student characteristics might have played a secondary role in the retention decision.

<sup>2</sup> Students who had been retained at the end of Grade 5 or 6 (after Wave 1 or 2) were not included in the analyses. Those students were no longer surveyed along with their classmates, but were followed individually.

panel wave, which occurred at the beginning of seventh grade (November 2012 to January 2013; Wave 3) and at the beginning of eighth grade (November 2013 to February 2014; Wave 4). In Wave 3 (January 2013), students were on average 12.5 years old ( $SD = 0.62$ ), 49.4% were female.

### Measures

The following measures were applied in data collection.

#### Retention or promotion status

We retrieved the information on whether a student had been retained or promoted from school records indicating the students' current grade level (Wave 3 in Sample 1; Wave 4 in Sample 2).

#### Personality characteristics

Students' personality traits were assessed using the 10-item version (BFI-10; Rammstedt & John, 2007) of the Big Five Inventory (John, Donahue, & Kentle, 1991), which has been shown to be a valid short-scale, applicable as an alternative to the longer original instrument (Wave 1 in Sample 1; Wave 3 in Sample 2). The BFI-10 comprises two items (one positively poled and one negatively poled) from each of the original scales of the Big Five Inventory. It assesses the extent to which a person is "outgoing, sociable" vs. "reserved" (Extraversion); "tends to be lazy" vs. "does a thorough job" (Conscientiousness); is "relaxed, handles stress well" vs. "gets nervous easily" (Emotional Stability); has "an active imagination" vs. "few artistic interests" (Openness to Experience); is "generally trusting" vs. "tends to find fault with others" (Agreeableness). The BFI-10 has been demonstrated to exhibit reasonable stabilities (test-retest reliabilities; Gosling, Rentfrow, & Swann Jr., 2003).<sup>3</sup> Students responded to each item on a 5-point scale (1 = *does not apply at all* to 5 = *fully applies*).

#### Teacher-assigned grades

We obtained teacher-assigned grades in mathematics and German from student self-reports. In Sample 1, students reported their mid-year grades for Year 9 in Wave 2. In Sample 2, students reported their final Year 7 grades in Wave 4. In the German school system, teachers assign numeric grades from 1 to 6, with a grade of 1 reflecting excellent achievement, and a grade of 6 denoting unsatisfactory achievement. For the analyses, we reverse-coded the grades so that a grade of 6 indicated excellent achievement, and a grade of 1 reflected unsatisfactory achievement.

#### Test scores

In Sample 1, test scores for mathematics and reading speed were available (Wave 1). The standardized achievement test for mathematics used in the NEPS incorporates 22 items from the content areas *quantity* (seven items), *space and shape* (six items), *change and relationships* (six items), and *data and chance* (three items), and comprises different cognitive components (*arguing, communicating, modeling, problem solving, representing, and applying technical skills*) that were distributed across the items (Durchhardt & Gerdes, 2013).<sup>4</sup> The response format was multiple choice (and short constructed response in the case of one item). Despite the different content areas, a unidimensional partial credit model was found to describe the data well (Durchhardt & Gerdes,

<sup>3</sup> For Sample 2, where the BFI-10 was administered twice—once in Grade 7 and once in Grade 9—retest-correlations were  $r_{tt} = 0.52$  for openness,  $r_{tt} = 0.51$  for conscientiousness,  $r_{tt} = 0.49$  for extraversion, and  $r_{tt} = 0.38$  for agreeableness and emotional stability, indicating a satisfactory reliability over the course of two years.

<sup>4</sup> Thus, the test is designed as a combination of the corresponding components from the German national educational standards for mathematics and PISA (Neumann et al., 2013).

2013). The WLE reliability was high (0.79; [Durchhardt & Gerdes, 2013](#)). In addition, the NEPS measured students' reading speed<sup>5</sup> using 51 items that were administered as a speed test: Students had to evaluate short sentences as either true or false (e.g., "There is a bathtub in every garage"). The items were based on the principles of the Salzburger Reading Screening which has been shown to be a reliable measure ( $r_{tt} = 0.89$ ; [Auer, Gruber, Mayringer, & Wimmer, 2005](#)). The measure mainly assesses reading processes, such as decoding, and therefore captures automatized reading ([Zimmermann, Gehrler, Artelt, & Weinert, 2012](#)), which predicts differences in reading comprehension ([Rosebrock & Nix, 2006](#)).

In Sample 2, test scores for mathematics and orthography were available (Wave 3). The standardized achievement test for mathematics used in Sample 2 included 23 items from the content areas *quantity* (five items); *space and shape* (five items); *change and relationships* (seven items); and *data and chance* (six items) and tapped different cognitive components (*arguing, communicating, modeling, problem solving, representing, and applying technical skills*) distributed across the items ([Schnittjer & Gerken, 2017](#))<sup>3</sup>. The items had a multiple-choice response format (with the exception of one item requiring a short constructed response). A unidimensional partial credit model yielded a good model fit ([Schnittjer & Gerken, 2017](#)). The WLE reliability was high (0.72; [Schnittjer & Gerken, 2017](#)). Furthermore, orthography<sup>4</sup> was measured using 137 items (119 words in full sentences and 18 words in cloze test) assessing five orthographic subskills (*phonographic syllabic subskill, morphological subskill, peripheral subskill, derivational subskill, and syntactic subskill*). Based on these items, a generalized score was computed. The reliability was high (0.95; [Blatt, Frahm, Jarsinski, & Prosch, 2014](#)).

#### Family background

Students reported their family backgrounds in both Wave 1 (Sample 1) and Wave 3 (Sample 2). Based on these reports, the International Socio-Economic Index of Occupational Status (ISEI-08; see [Ganzeboom, 2010](#)) was determined. The ISEI-08 has a theoretical range from 11.74 (low SES; e.g., manual workers in agricultural sectors who lack schooling) to 88.96 (high SES; e.g., judges). We used the higher value of the two parents (HISEI) in the analyses.

#### Days absent

Students reported the number of days they had been absent in the current school year (Wave 1 in Sample 1; Wave 3 in Sample 2). In Sample 1, students were absent for 3 days on average, with a range of 0 to 26 days. In Sample 2, the average number of days absent was 1 and ranged from 0 to 17 days.

#### Average test scores and school track (class-level variables)

At the class level, the analyses involved the average classrooms' test score across both test domains (in other words, we determined the mean value of the class-level aggregated standardized test scores for mathematics and German). We further included an indicator of the type of school track, as retention rates vary across school tracks ([German Federal Statistical Office, 2017](#)). The German system of secondary school is divided into different tracks. Only the academic track ("Gymnasium") leads to a university entrance certificate, while the vocational track ("Hauptschule") and the intermediate track ("Realschule") provide a more basic level of general education. The comprehensive school track ("Gesamtschule") combines these three tracks and exists parallel to them. The enrollment criteria for the academic track are much stricter (e.g., GPA at the end of primary school) than for the other tracks. Thus, in our analyses we differentiated between the academic track and all other tracks.

<sup>5</sup> The measure corresponds to one of the competencies outlined by German national educational standards (namely, listening, reading, spelling, writing, and language use; Institute for Educational Quality Improvement, n.d.).

#### Statistical analyses

To control for acquiescence ("yeah-saying"), we ipsatized items following the approach described by [John, Naumann, and Soto \(2008\)](#). Thus, we computed within-person average scores and standard deviations across all raw items and subsequently within-person standardized all items (see [Ashton et al., 2004](#); [Spengler et al., 2013](#)). Acquiescence seems to be more pronounced among lower-achieving individuals ([Lechner & Rammstedt, 2015](#)). It is of particular relevance for short scales such as the BFI-10, which captures each trait with one positively and one negatively formulated item only ([Rammstedt & John, 2007](#)). To investigate the factor structure, we conducted principal component analyses with an oblimin rotation in SPSS 24 (for a similar procedure, see [Spengler et al., 2013](#)). The eigenvalues supported the extraction of five factors with eigenvalues > 1 (Sample 1: 1.83, 1.61, 1.38, 1.23, 1.15, 0.83; Sample 2: 1.76, 1.59, 1.36, 1.13, 1.05, 0.88). The extracted variance of the five factors ranged from 69% (Sample 2) to 72% (Sample 1). The factor structure corresponded to the theoretically assumed Five Factor model, with each item exhibiting the highest loading on the corresponding factor (see Appendix). Factor correlations are shown in the Appendix.

For our main analyses, we computed multilevel logistic regression analyses (Mplus 7.4) to estimate the effects of student personality traits, test scores, SES, gender and days absent (Level 1, student level), along with the average classroom achievement and school type (Level 2, classroom level). Initially, we predicted grade promotion vs. retention from student personality only (Model 1, [Table 2](#)). In a next step, we only included teacher-assigned grades as predictors (Model 2, [Table 2](#)). We then specified a mediation model following the guidelines of [Zhang, Zyphur, and Preacher \(2009\)](#), with teacher-assigned grades as mediator variables and students' Big Five personality traits as independent variables (Model 3, [Table 3](#)). We included *test scores, socioeconomic background, gender and days absent* (Level 1, student level) as control variables and estimated their direct and indirect effects (via teacher-assigned grades). Moreover, we controlled for the classrooms' average test scores and school track, for which we also estimated direct and indirect effects (Level 2, classroom level). We estimated the indirect effect for each predictor by multiplying the regression coefficients of the outcome on grades (mediator) with the regression coefficients of grades on each specific predictor—separately for mathematics and German—and adding together the two product terms. All continuous student-level predictors were z-standardized and subsequently centered around their group mean. The average class-level test scores were aggregated first and subsequently z-standardized. The average percentages of missing values were 6.7% in Sample 1 and 8.6% in Sample 2. We applied the full-information maximum-likelihood approach (FIML; [Enders, 2001](#)) to acquire appropriate estimates and standard errors.

#### Results

[Fig. 1](#) depicts mean scores of all study variables separately for retained and promoted students in Sample 1 (Grade 9) and Sample 2 (Grade 7). The figure shows that the most substantial differences between retained and promoted students could be found for teacher-assigned grades (standardized differences > 0.90), while differences in standardized test scores were somewhat less pronounced (standardized differences ranged from 0.23 to 0.70). In addition, retained students had a lower SES (standardized differences of 0.28 in Sample 1 and 0.47 in Sample 2); had been absent for a longer period of time (standardized differences of 0.50 in Sample 1 and 0.25 in Sample 2); and scored lower in conscientiousness (standardized differences of 0.34 in Sample 1 and 0.36 in Sample 2) than promoted students. Standardized differences in the other four personality traits, i.e., openness, extraversion, agreeableness, and emotional stability, were smaller (0.16 or smaller). Correlations between all continuous study variables are presented in [Table 1](#).

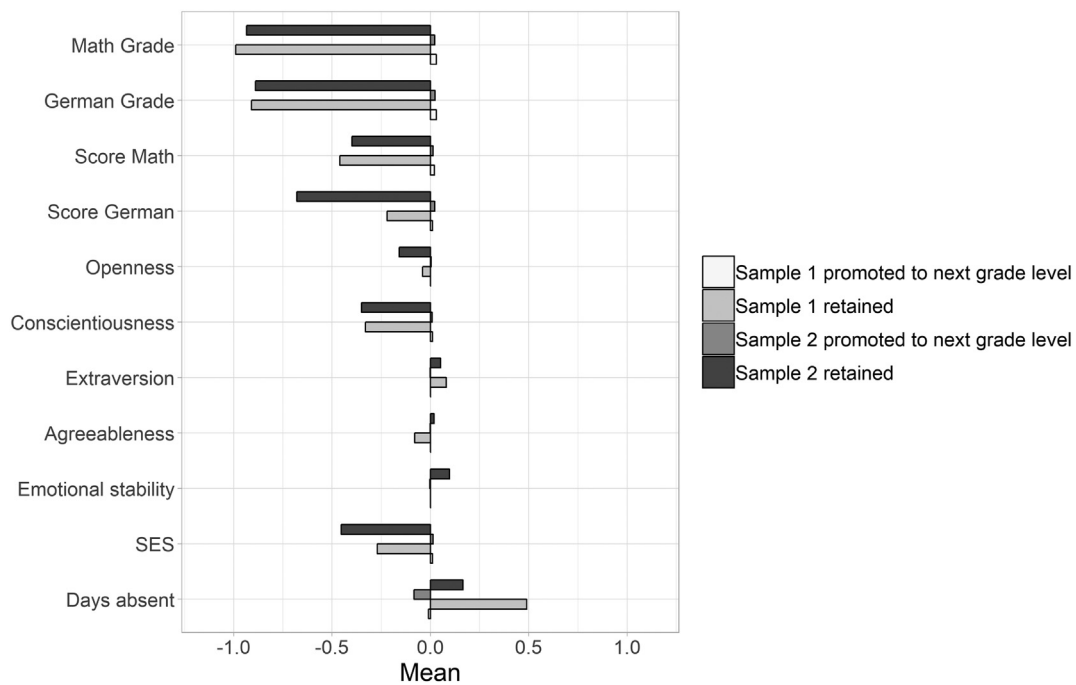


Fig. 1. Mean values of standardized study variables for retained and promoted students in Sample 1 (Grade 9) and Sample 2 (Grade 7).

Table 1  
Correlations of study variables in sample 1 (grade 9) and sample 2 (grade 7).

	1	2	3	4	5	6	7	8	9	10	11
1. Math grade		0.48***	0.41***	0.23***	0.01	0.15***	0.01	-0.01	0.09***	0.21***	-0.08***
2. German grade	0.38***		0.29***	0.47***	0.10***	0.19***	0.08***	-0.01	-0.02	0.23***	-0.10***
3. Score math	0.37***	0.21***		0.51***	0.10	-0.01	0.03	-0.07***	0.11***	0.35***	-0.10***
4. Score German	0.06***	0.22***	0.24***		0.11***	0.13***	0.08***	-0.01	0.02	0.30***	-0.13***
5. Openness	-0.02	0.12***	0.01	0.01		0.05***	0.02	0.07***	-0.03*	0.08***	0.02
6. Conscientiousness	0.16***	0.18***	-0.08***	0.00	0.06***		-0.03*	0.20***	0.04**	0.03	-0.09***
7. Extraversion	-0.06***	0.07***	-0.04***	0.09***	0.03**	-0.01		-0.10***	0.19***	0.08***	0.01
8. Agreeableness	0.00	-0.04***	-0.08***	-0.08***	0.02*	0.12***	-0.09***		-0.01	-0.01	-0.01
9. Emotional stability	0.05***	-0.01	0.10***	0.03**	-0.04***	-0.04***	0.20***	-0.03**		0.05**	-0.03*
10. SES	0.09***	0.15***	0.28***	0.17***	0.08***	-0.07***	0.06**	-0.05***	0.06***		-0.07**
11. Days absent	-0.13***	-0.06***	-0.11***	-0.04**	0.06***	-0.09***	0.03**	-0.04***	-0.02	0.00	

Correlations for sample 1 are depicted in the lower triangle, correlations for sample 2 in the upper triangle.

SES = Socioeconomic status.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

*Predicting grade retention from students' Big Five personality traits and teacher-assigned grades*

In order to estimate the effects of students' Big Five personality traits on grade retention, we specified a multilevel logistic regression model in which we predicted whether or not students were retained using all five personality traits as predictors (see Table 2, Model 1). The effects of students' conscientiousness on grade retention were statistically significant in both samples. Thus, less conscientious students were more likely to be retained.

To examine the effects of teacher-assigned grades on grade retention, we estimated a second multilevel logistic regression model predicting whether or not students were retained by teacher-assigned grades in mathematics and German (see Table 2, Model 2). We found that, in both samples, teacher-assigned grades were statistically significant predictors of grade retention. More specifically, ninth-grade students whose mathematics grades were one standard deviation below the average had 2.3 times the odds of being retained compared to students with average mathematics grades (while controlling for grades

in German). The effect sizes were similar for German grades and for Sample 2.

*Estimating direct and indirect effects of students' Big Five personality traits on grade retention (via teacher-assigned grades)*

Next, we tested: (a) to which extent students' Big Five personality traits affected teacher-assigned grades—which, in turn, had an impact on grade retention, and (b) to which extent students' Big Five personality traits affected grade retention above and beyond teacher-assigned grades (Table 3, Fig. 2). We also incorporated students' standardized test scores, gender, SES and the period of time they were absent from school as control variables. Moreover, we estimated the effect of the classrooms' average achievement and the school type on grade retention. First, the statistically significant relations between teacher-assigned grades and grade retention remained stable, with the exception of teacher-assigned grades in German in Sample 2, which no longer had a statistically significant effect on grade retention. In addition, students' standardized test scores in mathematics had statistically significant indirect effects

**Table 2**  
Multilevel logistic regression predicting grade retention (0 = proceeding to next grade, 1 = retention).

	Sample 2 (grades 7–8)															
	Model 1					Model 2										
	Est.	SE	OR	95% CI	Est.	SE	OR	95% CI	Est.	SE	OR	95% CI				
Intercept	3.79***	0.07**	44.30*	[38.44, 51.01]	4.41***	0.11	82.35	[67.02, 101.19]	4.27***	0.11	71.45	[57.11, 89.39]	4.86***	0.17	129.15	[91.84, 181.64]
Level 1																
Openness	0.00	0.06	1.00	[0.89, 1.11]					-0.12	0.10	0.89	[0.73, 1.07]				
Conscientious.	-0.37***	0.06	0.69	[0.62, 0.77]					-0.38***	0.10	0.69	[0.57, 0.83]				
Extraversion	0.11	0.06	1.12	[1.00, 1.25]					0.06	0.10	1.06	[0.87, 1.28]				
Agreeableness	-0.03	0.06	0.97	[0.87, 1.08]					0.11	0.10	1.11	[0.92, 1.35]				
Emotional stability	-0.04	0.06	0.96	[0.86, 1.07]					0.11	0.10	1.12	[0.92, 1.36]				
Math Grade					-0.85***	0.08	0.43	[0.37, 0.49]					-0.65***	0.12	0.52	[0.41, 0.66]
German Grade					-0.68***	0.07	0.51	[0.44, 0.59]					-0.53***	0.12	0.59	[0.47, 0.74]

Conscientious. = conscientiousness.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

on grade retention. Thus, students with lower standardized test scores in mathematics received lower teacher-assigned grades and consequently had a higher risk of being retained. The indirect effects of standardized scores in German were statistically significant in Sample 1, but not in Sample 2. We found direct effects on grade retention for standardized test scores in mathematics (Sample 1) and German (Sample 2). Thus, students who scored lower on standardized tests in German orthography and mathematics, respectively, had a higher risk of being retained than students with similar teacher-assigned grades, but higher standardized scores.

Concerning our third hypothesis, we found that students' conscientiousness had a statistically significant indirect effect on grade retention in both samples and a statistically significant direct effect in Sample 1. Thus, in both samples, less conscientious students exhibited lower teacher-assigned grades which, in turn, increased their risk of being retained. Moreover, less conscientious ninth-grade students were more likely to be retained than more conscientious students, even when they had similar teacher-assigned grades, standardized test scores, other personality traits and background characteristics. We also found a statistically significant indirect effect on grade retention (via teacher-assigned grades) of students' *agreeableness*, in the case of Sample 1, and students' *openness* and *emotional stability*, in the case of Sample 2. However, these effect sizes were extremely small.

Furthermore, in Sample 1, students' *SES*, *gender* and the *number of days they had been absent from school* were statistically significant indirect and direct predictors of grade retention. The indirect effects indicated that students who: were boys, had lower SES, and a higher number of days absent received lower grades which, in turn, affected whether or not they were retained. However, the effect sizes of students' *SES* and *days absent* were small. The direct effects indicate that students with lower SES were more likely to be retained than students who did not differ in teacher-assigned grades, standardized test scores, and personality, but had a higher SES. Similarly, boys were more likely to be retained than girls with similar grades, test scores, personality and background characteristics. Students who had been absent for a longer period of time were also more likely to be retained (holding all other characteristics equal). In Sample 2, we found similar direct effects. Concerning indirect effects in Sample 2, the seventh-grade students' *SES* and the length of time that they had been absent had an indirect effect on grade retention, while there was no significant indirect effect of students' gender.

Finally, in both samples, we found direct effects of average grades in mathematics and indirect effects of average test scores. Thus, students in classes with a higher average achievement received lower average grades and were therefore more likely to be retained. Moreover, in Sample 1, we found a statistically significant direct effect of the academic track on grade retention. In Sample 2, the indirect effect of the academic track on grade retention was statistically significant. Thus, controlling for average classroom achievement, ninth-grade students in academic track schools were less likely to be retained than students in other tracks. In contrast, seventh-grade students in academic-track schools received lower average grades and were therefore more likely to be retained.

## Discussion

Retaining students for one school year is a measure thought to enable low-achieving students to catch up with the curriculum, although its long-term effectiveness has been questioned (Allen et al., 2009). In Germany and similarly in the United States, the percentage of students who are retained lies in the mid-range of all OECD countries (Ikeda & García, 2014). Nevertheless, one-fifth of all 15-year-old students in Germany have been retained at least once in their previous school career (Ikeda & García, 2014), with adolescent students being retained most frequently (German Federal Statistical Office, 2017). The present study examined a factor that has repeatedly been linked to adolescents'

**Table 3**  
Multilevel mediation model predicting grade retention (0 = proceeding to next grade, 1 = retention).

	Sample 1 (grades 9–10)				Sample 2 (grades 7–8)				
	Model 3				Model 3				
	Est.	SE	OR	95% CI	Est.	SE	OR	95% CI	
Intercept	4.34***	0.13	76.71	[59.74, 97.51]	5.19***	0.23	179.47	[113.3, 281.46]	
Level 1 (direct effects)									
Math grade	−0.71***	0.08	0.49	[0.42, 0.58]	−0.70***	0.14	0.50	[0.38, 0.65]	
German grade	−0.66***	0.08	0.52	[0.44, 0.61]	−0.01	0.14	0.99	[0.75, 1.30]	
Score math	−0.38***	0.10	0.68	[0.56, 0.83]	0.01	0.15	0.99	[0.73, 1.33]	
Score German	−0.08	0.07	0.92	[0.81, 1.06]	−0.84***	0.16	0.43	[0.32, 0.59]	
Openness	0.04	0.06	1.04	[0.92, 1.18]	0.03	0.11	1.03	[0.82, 1.28]	
Conscientious.	−0.15*	0.07	0.86	[0.75, 0.98]	−0.10	0.11	0.91	[0.73, 1.13]	
Extraversion	0.12	0.06	1.13	[0.99, 1.28]	0.10	0.11	1.11	[0.89, 1.38]	
Agreeableness	−0.09	0.06	0.91	[0.81, 1.02]	0.07	0.11	1.07	[0.86, 1.33]	
Emotional stab.	0.06	0.06	1.06	[0.93, 1.20]	0.07	0.11	1.08	[0.86, 1.34]	
SES	−0.25**	0.08	0.78	[0.67, 0.91]	−0.34*	0.16	0.72	[0.53, 0.97]	
Gender <sup>a</sup>	0.35**	0.13	1.42	[1.10, 1.82]	0.98***	0.23	2.68	[1.71, 4.19]	
Days absent	0.29***	0.06	1.33	[1.20, 1.48]	0.30**	0.09	1.35	[1.13, 1.61]	
Level 1 (indirect effects)									
Score math	−0.50***	0.06	0.61	[0.54, 0.68]	−0.31***	0.06	0.73	[0.66, 0.82]	
Score German	−0.12***	0.02	0.89	[0.86, 0.91]	−0.07	0.07	0.93	[0.83, 1.06]	
Openness	−0.02	0.01	0.98	[0.96, 1.01]	0.02*	0.01	1.02	[1.00, 1.04]	
Conscientious.	−0.24***	0.02	0.79	[0.76, 0.81]	−0.11***	0.02	0.90	[0.86, 0.94]	
Extraversion	−0.01	0.01	0.99	[0.97, 1.02]	0.01	0.01	1.01	[0.98, 1.03]	
Agreeableness	0.02**	0.01	1.02	[1.01, 1.04]	0.01	0.01	1.01	[0.99, 1.03]	
Emotional stab.	0.00	0.01	1.00	[0.98, 1.01]	−0.03**	0.01	0.97	[0.95, 0.99]	
SES	−0.10***	0.01	0.90	[0.89, 0.93]	−0.07***	0.02	0.93	[0.89, 0.97]	
Gender <sup>a</sup>	0.12***	0.02	1.13	[1.09, 1.17]	−0.03	0.02	0.97	[0.94, 1.01]	
Days absent	0.10***	0.01	1.11	[1.08, 1.14]	0.03**	0.01	1.03	[1.01, 1.05]	
Level 2 (direct effects)									
Avg. math grade	−0.60***	0.13	0.55	[0.43, 0.70]	−0.57**	0.20	0.57	[0.39, 0.84]	
Avg. Ger. grade	−0.14	0.13	0.87	[0.68, 1.13]	−0.36	0.23	0.70	[0.44, 1.09]	
Avg. test score	0.03	0.11	1.03	[0.83, 1.28]	0.11	0.18	1.12	[0.79, 1.57]	
Academic track <sup>b</sup>	−0.73***	0.22	0.48	[0.31, 0.75]	−0.55	0.29	0.58	[0.32, 1.02]	
Level 2 (indirect effects)									
Avg. test score	−0.10***	0.03	0.90	[0.87, 0.95]	−0.22***	0.06	0.80	[0.71, 0.91]	
Academic track <sup>b</sup>	0.03	0.04	1.03	[0.95, 1.11]	−0.10*	0.05	0.90	[0.82, 1.00]	

Conscientious. = conscientiousness. SES = Socioeconomic status. Avg. math grade = average math grade. Avg. Ger. grade = average German grade. Avg. test score = average test score.

<sup>a</sup> Reference: female.

<sup>b</sup> Reference: all other tracks.

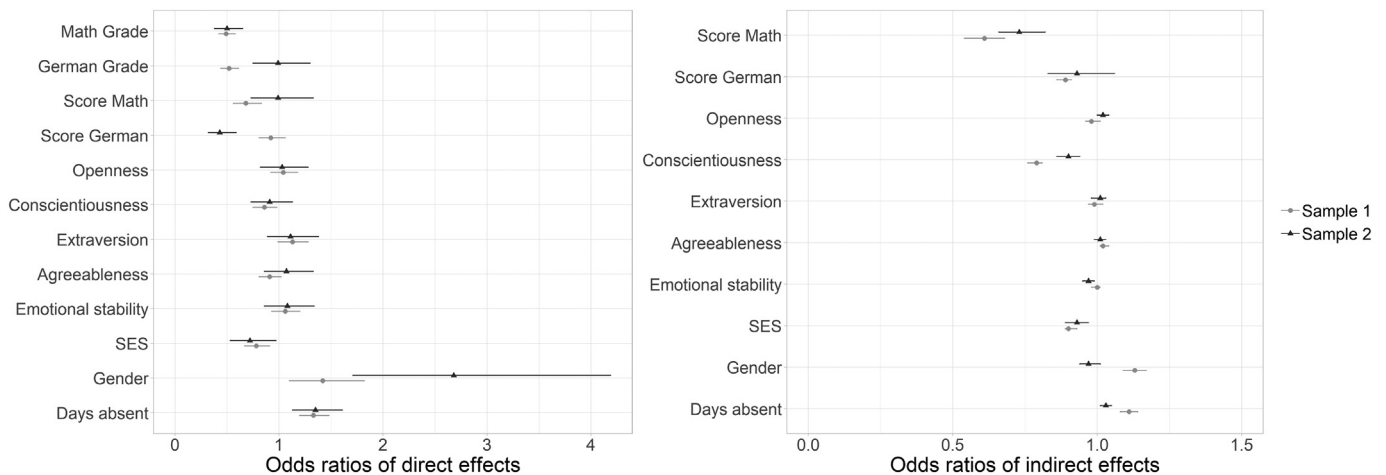
\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

academic success (Spengler et al., 2013), namely the role of students' Big Five personality traits, in grade retention. Moreover, we aimed at clarifying the potential mechanisms responsible for any relationships between student personality and grade retention.

Initially we found that, in both samples, less conscientious students were more likely to be retained, while the other personality traits were not statistically significant predictors of grade retention. This finding confirmed Hypothesis 1 and is consistent with previous research (Laidra



**Fig. 2.** Odds ratios of direct and indirect effects of student-level variables on grade retention in Sample 1 (Grades 9–10) and Sample 2 (Grades 7–8).



et al., 2007; Spengler et al., 2013) indicating that conscientiousness may be a very good predictor of academic outcomes, particularly in secondary school. Our initial models also showed that teachers' decisions to retain a student were closely associated with teacher-assigned grades, thus, confirming Hypothesis 2. This is in line with state laws and regulations in Germany, which impose teacher-assigned grades as major criteria for the decision to retain a student.

Our mediation analyses also revealed that teachers' decision to retain a student was first and foremost based on students' grades. However, seventh-grade students' standardized test scores in orthography were better predictors for grade retention than teacher-assigned grades in German. This finding could potentially be explained by the fact that our study lacks teacher-assigned grades in subjects other than mathematics and German—which, however, may also play a role in the decision to retain or promote a student. It is conceivable that teacher-assigned grades, especially in liberal arts and languages, may be more closely tied to students' standardized test scores in orthography than to their teacher-assigned grades in German. Thus, students' orthography scores may be a proxy for teacher-assigned grades in other subjects and therefore are a strong predictor of students' retention.

We also found indirect effects of students' scores in standardized mathematics tests on grade retention. Moreover, ninth-grade students' scores in mathematics and seventh-grade students' scores in orthography exhibited direct effects on grade retention. That students with lower scores in standardized tests in mathematics received less favorable grades and therefore had a higher risk of being retained likely reflects the fact that teacher-assigned grades are linked to standardized test scores (Westphal et al., 2016; see also Wentzel, Weinberger, Ford, & Feldman, 1990; but also see Valtin, Badel, Löffler, Meyer-Schepers, & Voss, 2003 for findings on deviations between teacher-assigned grades and standardized test scores). The direct effects of the standardized test scores in orthography in Sample 2 and mathematics in Sample 1 might indicate that these scores are good proxies for teacher-assigned grades in other subjects that were not incorporated in our study.

#### *What are the processes by which the Big Five personality traits in students affect grade retention?*

We found that students' conscientiousness had an indirect effect on teachers' retention decision via teacher-assigned grades. In line with Hypothesis 3a, less conscientious students therefore received less favorable grades which, in turn, increased their risk of being retained. Conscientiousness in adolescents may therefore be highly relevant in their disengagement from schooling—seen as one of the main challenges for educators (Wigfield et al., 2006)—which in turn has been shown to affect teacher-assigned grades (Wang & Holcombe, 2010). Thus, lower levels of self-control and academic motivation in less conscientious adolescents (Lounsbury et al., 2004; MacCann et al., 2009) could affect the time they spend learning (for a similar suggestion see Poropat, 2014). These findings corroborate previous research (Huang, 2014; Martin, 2009) suggesting that factors such as task persistence and homework completion predict grade retention, and correspond to previous studies (Laidra et al., 2007; Spengler et al., 2013) showing that less conscientious adolescents receive lower teacher-assigned grades. Furthermore, we found statistically significant indirect effects on grade retention of agreeableness in ninth-grade students and openness and emotional stability in seventh-grade students. These effects were small, however, and should consequently be interpreted with caution.

Our findings also partially confirmed Hypothesis 3b by showing that ninth-grade students' conscientiousness predicted grade retention over and above teacher-assigned grades and a broad range of student demographics. Thus, ninth-grade students with similar teacher-assigned grades had a higher risk of being retained if they reported a lower degree of conscientiousness. This result adds to reports (Bless et al., 2004; Tomchin & Impara, 1992; Witmer et al., 2004) indicating that

teachers take student's effort and motivation into account when deciding on a student's retention. The reasons why teachers consider ninth-grade students' conscientiousness in their retention decisions may be twofold. On the one hand, teachers may aim at maximizing a student's chance of being successful in the next grade level (Tomchin & Impara, 1992; see also Range, Pijanowski, Holt, & Young, 2012), which may be seen as more likely if low-achieving students exhibit higher levels of conscientiousness. On the other hand, teachers may also apply grade retention in order to sanction students with unfavorable work habits (Tomchin & Impara, 1992; Wong & Zhou, 2017), thereby aiming at increasing parents' and students' effort put into schoolwork (Range et al., 2012; Wong & Zhou, 2017). The direct effect of conscientiousness on grade retention was not significant, however, in the sample of seventh-grade students. It may be that teachers base their retention decisions on different indicators in Grade 9 compared to Grade 7. Teachers of ninth-grade students may consider the fact that, in Grade 10, nonacademic-track students face important challenges, as they have to apply for an apprenticeship position (which requires favorable grades and much effort when searching for potential positions and preparing applications). Therefore, teachers may place more weight on students' conscientiousness when retaining or promoting low-achieving ninth-grade students than when deciding upon seventh-grade students' retention. This would also be in line with the results of the qualitative study by Tomchin and Impara (1992), in which teachers reported that noncognitive factors are differentially relevant in different grade levels.

Apart from conscientiousness, none of the other Big Five personality traits had a statistically significant direct effect on grade retention. This reflects qualitative findings (Tomchin & Impara, 1992; Wong & Zhou, 2017) in which teachers predominantly named different aspects of conscientiousness, such as self-discipline, order, achievement striving, but hardly mentioned any aspects of the other personality traits as relevant criteria for retention decisions.

#### *Do SES, gender, absence and class-level achievement additionally explain grade retention?*

In addition, SES had a direct effect and a small indirect effect on grade retention. While the indirect effect was modest and should be interpreted with caution, the direct effect reveals that lower-SES students have a higher risk of being retained than higher-SES students, even when they exhibit similar teacher-assigned grades and personality traits. These results potentially reflect that teachers believe that low-achieving lower-SES students receive less parental support and therefore are less likely to succeed in the next grade level than low-achieving higher-SES students (for the relevance of parental support for teachers' retention decision, see Witmer et al., 2004). On the other hand, our findings could also be attributed to discrepancies in teacher judgments of students from varying socioeconomic backgrounds. This would be consistent with studies showing that teachers evaluate students' achievement less favorably if they perceive a mismatch between their own educational values and those of the student's parents (Hauser-Cram, Sirin, & Stipek, 2003; see also Sirin, Ryce, & Mir, 2009) or if they see a mismatch between the school's behavioral norms and the conduct of the child (Entwisle et al., 2007)—an issue that is more likely to arise for lower-SES students (Stephens, Markus, & Phillips, 2014). These socioeconomic disparities in students' risk of being retained are delicate, considering governmental aspirations to decouple educational attainment from family background.

Students' gender also had a direct effect on grade retention, in that boys were more likely to be retained than girls with similar teacher-assigned grades, personality traits and SES. In Sample 1, gender also had an indirect effect on grade retention, indicating that boys receive lower grades than girls with similar standardized achievement and personality traits, and are therefore more likely to be retained. Both effects may reflect teachers' tendency to judge boys less favorably than girls—a phenomenon that has been linked to more conduct problems in

males (e.g., Entwisle et al., 2007) and, relatedly, teachers' gender-stereotypic expectations (Kuhl & Hannover, 2012). Furthermore, the number of school days missed had a direct effect on grade retention, indicating that teachers were more likely to retain students who were absent more often, despite similar grades and personality traits (as were reported to play a role by at least some teachers in prior studies; Witmer et al., 2004).

On the class level, the average achievement in class had a positive indirect effect on grade retention, reflecting the well-documented “Big-Fish-Little-Pond Effect,” which denotes the frame-of-reference effect, i.e., that a student will receive less favorable teacher judgments in higher-achieving classes than in lower-achieving classes (Dompnier, Pansu, & Bressoux, 2006). In Sample 1, we also found a direct effect of school track, suggesting that students in the academic track are less likely to be retained than students in similar-achieving classes in other school tracks. This adds to previous findings (e.g., Ehmke et al., 2008) showing that retention is less frequent in the academic-track.

### Limitations and future research

Several limitations of this study should be mentioned. First, it has been suggested (Spengler et al., 2013) that when personality is measured with the short BFI-10 scale using two items per factor, variance in personality traits may be decreased and, as a consequence, the role of personality in academic outcomes could be underestimated (Credé, Harms, Niehorster, & Gaye-Valentine, 2012). Yet, there is evidence that the BFI-10 scale, in fact, measures 70% of the variance captured by the full BFI (Rammstedt & John, 2007). Moreover, relationships between academic outcomes and personality traits have been found to be similar when measured with the BFI-10 scale, as compared to longer personality inventories (Spengler et al., 2013). Nevertheless, we cannot rule out that our results only reveal the lower-bound relationships between students' personality traits and teachers' retention decisions.

Secondly, we did not have teacher-assigned grades for all subjects available in our study. Because grades in mathematics and German are crucial in a teacher's decision to retain a student, we focused on grades in these subject domains in our study. However, depending on the school track and federal state, teachers may also take grades in other subjects into account, for instance in foreign languages. Thus, direct effects of students' conscientiousness on grade retention may, at least partially, reflect that less conscientious students receive lower grades in other subjects as well and are therefore at higher risk of being retained. In addition, we used students' self-reported grades, which may depend on students' level of achievement (Kuncel, Credé, & Thomas, 2005). Therefore, we might be underestimating the actual effects of teacher-assigned grades on retention and, consequently, our indirect effects, and we might be overestimating the effects of standardized tests on grade retention. Furthermore, our analyses only included students for whom there were available data on their retention status. Data on retention status were missing in the case of a number of students for several reasons, including, for example, that schools opted out of the study or that students changed to other school forms (e.g., vocational schools). When we compared students for whom we had data about their retention status with those for whom we did not, we found that students for whom we did not have retention data had slightly lower socio-economic statuses and teacher-assigned grades. Thus, further studies are needed that particularly focus on students from lower socio-economic backgrounds and lower grades to examine whether the findings of our study can be replicated in these specific groups.

Thirdly, we did not find a direct effect of seventh-grade students' conscientiousness on grade retention, which may reflect differential grading strategies depending on students' age. However, these results could be attributed, in part, to the differences in the study variables in Sample 1 and Sample 2. For example, we had final-year grades available for seventh-grade students, but only mid-year grades for ninth-grade students, and slightly different standardized test scores.

Differences in the curricular validity of the standardized tests may therefore contribute to different relationships between test scores and grades in the two samples (Willingham, Pollack, & Lewis, 2002). Future research should incorporate teacher-assigned grades in a broad range of subjects. It may also be promising to include open-ended questions assessing the strategies teachers use when deciding on students' retention in these different points of students' educational career. Finally, studies incorporating Big Five personality traits in teachers may be enlightening, as previous research shows that students receive better grades if their personality is similar to that of their teachers (e.g., Rausch, Karing, Dörfler, & Artelt, 2016).

### Practical implications and conclusions

While previous studies on the role of students' Big Five personality traits (Laidra et al., 2007; Poropat, 2014; Spengler et al., 2013) focused on teacher-assigned grades showing that a student's conscientiousness may be crucial for their academic success, this study is the first to link students' conscientiousness to grade retention. Moreover, students' conscientiousness affected grade retention beyond a wide range of student characteristics and even showed incremental validity above teacher-assigned grades when explaining ninth-grade students' retention. Such direct effects would be particularly problematic if they indeed reflected teacher intentions to sanction unfavorable work habits, as suggested by qualitative research (Tomchin & Impara, 1992; Wong & Zhou, 2017). However, even if the teachers' intention—when retaining less conscientious students and promoting similar achieving, but more conscientious, students—is to maximize the students' chance of being successful in the next school year (as suggested by the results of Tomchin & Impara, 1992), the question remains whether such prognoses are in fact accurate. Given the economic costs—in the form of eluded earnings and per student spending that result if students remain an additional year in the educational system—especially false negative judgments should be avoided. Thus, teachers should be encouraged to use grade retention judiciously.

Furthermore, our findings indicate that boys and lower-SES students are more likely to be retained than girls and higher-SES students with similar teacher-assigned grades and personality characteristics. Thus, teacher training should raise teachers' awareness for their own potential liability to gender and socioeconomic bias. Even if teachers feel that lower-SES students receive less support and may therefore seem less likely to overcome difficulties in school (Witmer et al., 2004), retaining lower-SES students based on their background, instead of their achievement in school, runs counter to the distributive justice that is aspired by those educational systems underpinned by democratic and egalitarian educational ideals. Some states in the United States have introduced test-based, instead of teacher-based, retention (for an overview, see Huddleston, 2014). This may, on the one hand, help to reduce socioeconomic discrepancies in teacher judgments by grounding retention decisions more strongly on actual student achievement. On the other hand, test-based retention policies may itself disadvantage lower-SES or minority students as well as students in lower-quality schools (Livingston & Livingston, 2002) and have been criticized for “reinforcing the ideology that success on high-stakes tests is solely the result of effort while masking the connection between educational achievement and social inequities” (Huddleston, 2014, p. 21). Nevertheless, standardized tests may help teachers by directing their attention to relevant dimensions of academic development (Kowalski, Brown, & Pretti-Frontczak, 2005). Moreover, standardized tests offer feedback that teachers can use to verify their judgments (Helmke, Hosenfeld, & Schrader, 2004). As such, teachers can explicate the perceptions they have of specific students and seek explanations if a student's actual results deviate from the teacher's expectations (Helmke et al., 2004). Our study points to several aspects that, taken together, could help inform teachers and teacher trainers who seek to further improve decisions on grade retention.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.appdev.2019.101088>.

## References

- Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis, 31*, 480–499. <https://doi.org/10.3102/0162373709352239>.
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., ... De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives. Solutions from psychlexical studies in seven languages. *Journal of Personality and Social Psychology, 86*, 356–366.
- Auer, M., Gruber, G., Mayringer, H., & Wimmer, H. (2005). *Salzburger Lesescreening für die Klassenstufen 5–8 [Salzburger screening of reading competencies in grade levels 5–8]*. Göttingen, DE: Hogrefe.
- Bandura, A. (1964). The stormy decade: Fact or fiction? *Psychology in the Schools, 1*(3), 224–231.
- Blatt, I., Frahm, S., Jarsinski, S., & Prosch, A. (2014). *Technical report for orthography: Scaling results of starting cohort 3 in grades 5 and 7 (NEPS working paper no. 49)*. Bamberg, DE: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Bless, G., Schüpbach, M., & Bonvin, P. (2004). *Klassenwiederholung. Determinanten, Wirkungen und Konsequenzen [Grade retention: Predictors, effects and consequences]*. Bern: Haupt-Verlag.
- Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (Eds.). (2011). Education as a lifelong process—The German National Educational Panel Study (NEPS) [special issue], *Zeitschrift für Erziehungswissenschaft, 14*.
- Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the big five personality traits. *Journal of Personality and Social Psychology, 102*, 874–888. <https://doi.org/10.1037/a0027403>.
- Damian, R. L., Spengler, M., Sutu, A., & Roberts, B. W. (2018). Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspp0000210> Advance online publication.
- Dauber, S. L., Alexander, K. L., & Entwisle, D. R. (1993). Characteristics of retainees and early precursors of retention in grade: Who is held back? *Merrill-Palmer Quarterly, 39*, 326–343.
- Dompnier, B., Pansu, P., & Bressoux, P. (2006). An integrative model of scholastic judgments: Pupils' characteristics, class context, halo effect and internal attributions. *European Journal of Psychology of Education, 21*, 119–133. <https://doi.org/10.1007/BF03173572>.
- Durchhardt, C., & Gerdes, A. (2013). *NEPS technical report for mathematics—Scaling results of starting cohort 4 in ninth grade (NEPS working paper no. 22)*. Bamberg, DE: University of Bamberg, National Educational Panel Study.
- Ehmke, T., Drechsel, B., & Carstensen, C. H. (2008). Klassenwiederholen in PISA-I-Plus: Was lernen Sitzbleiber in Mathematik dazu [Grade retention in PISA-I-Plus: What do retained children learn in mathematics]? *Zeitschrift für Erziehungswissenschaft, 11*, 368–387. <https://doi.org/10.1007/s11618-008-0033-3>.
- Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods, 6*, 352–370. <https://doi.org/10.1037/1082-989X.6.4.352>.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education, 80*, 114–138. <https://doi.org/10.1177/003804070708000202>.
- Erikson, E. H. (1959). *Identity and the life cycle: Selected papers*. Oxford, England: International Universities Press.
- Ganzeboom, H. B. G. (2010, May). A new international socio-economic index (ISEI) of occupational status for the international standard classification of occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007; With an analysis of quality of occupational measurement in ISSP. Paper presented at the Annual Conference of International Social Survey Programme (ISSP), Lisbon, Portugal. Retrieved from <http://www.harryganzeboom.nl/pdf/2010-ganzeboom-isei08-issp-lisbon-paper.pdf>.
- German Federal Statistical Office (2004). Bildung und Kultur. Allgemeinbildende Schulen [Education and culture. General education schools]. Retrieved from [https://www.destatis.de/GPstatistik/servlets/MCRFileNodeServlet/DEHeft\\_derivate\\_00006810/2110100047004.pdf](https://www.destatis.de/GPstatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00006810/2110100047004.pdf).
- German Federal Statistical Office (2017). Bildung und Kultur. Allgemeinbildende Schulen [Education and culture. General education schools]. Retrieved from [https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100167004.pdf?\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100167004.pdf?_blob=publicationFile).
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality, 37*, 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1).
- Hauser-Cram, P., Sirin, S. R., & Stipek, D. (2003). When teachers' and parents' values differ: Teachers' ratings of academic competence in children from low-income families. *Journal of Educational Psychology, 95*, 813–820. <https://doi.org/10.1037/0022-0663.95.4.813>.
- Havighurst, R. J. (1948). *Developmental tasks and education*. New York: McKay.
- Helmke, A., Hosenfeld, I., & Schrader, F.-W. (2004). Vergleichsarbeiten als Werkzeug für die Verbesserung der diagnostischen Kompetenz von Lehrkräften [Comparison tests as a measure for improving the diagnostic competencies of teachers]. In R. Arnold, & C. Griesse (Eds.). *Schulleitung und Schulentwicklung [School management and school development]* (pp. 119–144). Hohengehren, DE: Schneider.
- Huang, F. L. (2014). Further understanding factors associated with grade retention: Birthday effects and socioemotional skills. *Journal of Applied Developmental Psychology, 35*, 79–93. <https://doi.org/10.1016/j.appdev.2013.12.004>.
- Huddleston, A. P. (2014). Achievement at whose expense? A literature review of test-based grade retention policies in U.S. schools. *Education Policy Analysis Archives, 22*(18), 1–31. <https://doi.org/10.14507/epaa.v22n18.2014>.
- Ikeda, M., & García, E. (2014). Grade repetition. A comparative study of academic and non-academic consequences. *OECD Journal: Economic Studies, 2013*(1), 269–315. [https://doi.org/10.1787/eco\\_studies-2013-5k3w65mx3hnx](https://doi.org/10.1787/eco_studies-2013-5k3w65mx3hnx).
- Im, M. H., Hughes, J. N., Kwok, O. M., Puckett, S., & Cerda, C. A. (2013). Effect of retention in elementary grades on transition to middle school. *Journal of School Psychology, 51*, 349–365. <https://doi.org/10.1016/j.jsp.2013.01.004>.
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30*, 420–437.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five inventory—versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.). *Handbook of personality: Theory and research* (pp. 114–158). (3rd ed.). New York, NY: Guilford.
- Klapproth, F., & Schaltz, P. (2015). Who is retained in school, and when? Survival analysis of predictors of grade retention in Luxembourgish secondary school. *European Journal of Psychology of Education, 30*, 119–136. <https://doi.org/10.1007/s10212-014-0232-7>.
- Kowalski, K., Brown, R. D., & Pretti-Frontczak, K. (2005). The effects of using formal assessment on preschool teachers' beliefs about the importance of various developmental skills and abilities. *Contemporary Educational Psychology, 30*, 23–42. <https://doi.org/10.1016/j.cedpsych.2004.05.001>.
- Kretschmann, J., Vock, M., Lüdtke, O., Jansen, M., & Gronostaj, A. (2017, August). How grade retention affects students' motivation: A longitudinal study over 3 years of secondary school. Presentation at the Conference of the European Association for Research on Learning and Instruction (EARLI), Tampere, Finland.
- Kühl, P., & Hannover, B. (2012). Differenzielle Benotungen von Mädchen und Jungen [Differential grading of girls and boys]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 44*, 153–162. <https://doi.org/10.1026/0049-8637/a000066>.
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63–82. <https://doi.org/10.3102/00346543075001063>.
- 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. <https://doi.org/10.1016/j.paid.2006.08.001>.
- Lechner, C. M., & Rammstedt, B. (2015). Cognitive ability, acquiescence, and the structure of personality in a sample of older adults. *Psychological Assessment, 27*, 1301–1311. <https://doi.org/10.1037/pas0000151>.
- Livingston, D. R., & Livingston, S. M. (2002). Failing Georgia: The case against the ban on social promotion. *Education Policy Analysis Archives, 10*(49), 1–31 (Retrieved from <https://epaa.asu.edu/ojs/article/view/328/454>).
- Lounsbury, J. W., Steel, R. P., Loveland, J. M., & Gibson, L. W. (2004). An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence, 33*(5), 457–466. <https://doi.org/10.1023/B:JOYO.0000037637.20329.97>.
- MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19*(4), 451–458. <https://doi.org/10.1016/j.lindif.2009.03.007>.
- Marsh, H. W. (2016). Cross-cultural generalizability of year in school effects: Negative effects of acceleration and positive effects of retention on academic self-concept. *Journal of Educational Psychology, 108*, 256–273. <https://doi.org/10.1037/edu0000059>.
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Lichtenfeld, S. (2017). Long-term positive effects of repeating a year in school: Six-year longitudinal study of self-beliefs, anxiety, social relations, school grades, and test scores. *Journal of Educational Psychology, 109*, 425–438. <https://doi.org/10.1037/edu0000144>.
- Martin, A. J. (2009). Age appropriateness and motivation, engagement, and performance in high school: Effects of age within cohort, grade retention, and delayed school entry. *Journal of Educational Psychology, 101*, 101–114. <https://doi.org/10.1037/a0013100>.
- Martin, A. J. (2011). Holding back and holding behind: Grade retention and students' non-academic and academic outcomes. *British Educational Research Journal, 37*, 739–763. <https://doi.org/10.1080/01411926.2010.490874>.
- Masten, A. S., Obradović, J., & Burt, K. B. (2006). Resilience in emerging adulthood: Developmental perspectives on continuity and transformation. In J. J. Arnett, & J. L. Tanner (Eds.). *Emerging adults in America: Coming of age in the 21st century* (pp. 173–190). Washington, DC: American Psychological Association.
- McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*, 273–298. [https://doi.org/10.1016/S0022-4405\(99\)00012-6](https://doi.org/10.1016/S0022-4405(99)00012-6).
- McGeown, S. P., Putwain, D., Simpson, E. G., Boffey, E., Markham, J., & Vince, A. (2014). Predictors of adolescents' academic motivation: Personality, self-efficacy and adolescents' characteristics. *Learning and Individual Differences, 32*, 278–286. <https://doi.org/10.1016/j.lindif.2014.05.001>.

- org/10.1016/j.lindif.2014.03.022.
- Meisels, S. J. (1998). *Assessing readiness (CIERA report #3-002)*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.
- 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/Journal für Bildungsforschung Online*, 5(2), 80–109.
- Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology*, 24(2), 124–139.
- Poropat, A. E. (2014). A meta-analysis of adult-rated child personality and academic performance in primary education. *British Journal of Educational Psychology*, 84, 239–252. <https://doi.org/10.1111/bjep.12019>.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>.
- Range, B. G., Pijanowski, J., Holt, C. R., & Young, S. (2012). The perceptions of primary grade teachers and elementary principals about the effectiveness of grade-level retention. *Professional Educator*, 36(1), n1.
- Rausch, T., Karing, C., Dörfler, T., & Artelt, C. (2016). Personality similarity between teachers and their students influences teacher judgement of student achievement. *Educational Psychology*, 36(5), 863–878. <https://doi.org/10.1080/01443410.2014.998629>.
- Rosebrock, C., & Nix, D. (2006). Forschungsüberblick: Leseflüssigkeit (Fluency) in der amerikanischen Leseforschung und –didaktik [Review: US-American research on reading fluency]. *Didaktik Deutsch*, 20, 90–112.
- Schnittjer, I., & Gerken, A.-L. (2017). *NEPS technical report for mathematics: Scaling results of starting cohort 3 in grade 7 (NEPS survey paper no. 16)*. Bamberg, DE: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- School Law of the Federal State Baden-Wuerttemberg GBl (2004). S. 397 § 1. Retrieved from <http://www.landesrecht-bw.de/jportal/?quelle=jlink&query=GymVersV+BW&psml=bsbawueprod.psml&max=true&aiz=true#jlr-GymVersVBWV7P1>.
- School Law of the Federal State North Rhine-Westphalia GV (2005). NRW. S. 102 § 50. Retrieved from <https://www.schulministerium.nrw.de/docs/Recht/Schulrecht/Schulgesetz/Schulgesetz.pdf>.
- Sirin, S. R., Rycie, P., & Mir, M. (2009). How teachers' values affect their evaluation of children of immigrants: Findings from Islamic and public schools. *Early Childhood Research Quarterly*, 24, 463–473. <https://doi.org/10.1016/j.ecresq.2009.07.003>.
- Spengler, M., Lüdtke, O., Martin, R., & Brunner, M. (2013). Personality is related to educational outcomes in late adolescence: Evidence from two large-scale achievement studies. *Journal of Research in Personality*, 47, 613–625. <https://doi.org/10.1016/j.jrp.2013.05.008>.
- 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. <https://doi.org/10.1016/j.paid.2009.11.028>.
- Steinmayr, R., & Spinath, B. (2007). Predicting school achievement from motivation and personality. *Zeitschrift für Pädagogische Psychologie*, 21(3/4), 207–216.
- Stephens, N. M., Markus, H. R., & Phillips, L. T. (2014). Social class culture cycles: How three gateway contexts shape selves and fuel inequality. *Annual Review of Psychology*, 65, 611–634. <https://doi.org/10.1146/annurev-psych-010213-115143>.
- Tanner, C. K., & Galis, S. A. (1997). Student retention: Why is there a gap between the majority of research findings and school practice? *Psychology in the Schools*, 34, 107–114. [https://doi.org/10.1002/\(SICI\)1520-6807\(199704\)34:2<107::AID-PITS4>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1520-6807(199704)34:2<107::AID-PITS4>3.0.CO;2-N).
- Tomchin, E. M., & Impara, J. C. (1992). Unraveling teachers' beliefs about grade retention. *American Educational Research Journal*, 29, 199–223. <https://doi.org/10.3102/00028312029001199>.
- Valtin, R., Badel, I., Löffler, I., Meyer-Schepers, U., & Voss, A. (2003). Orthographische Kompetenzen von Schülerinnen und Schülern der vierten Klasse [Orthographic competence of fourth-grade students]. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther, & R. Valtin (Eds.). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich [First IGLU results: Student achievement at the end of fourth grade in international comparison]* (pp. 227–264). Berlin, DE: Waxmann.
- Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, 47(3), 633–662. <https://doi.org/10.3102/0002831209361209>.
- Wentzel, K. R., Weinberger, D. A., Ford, M. E., & Feldman, S. S. (1990). Academic achievement in preadolescence: The role of motivational, affective, and self-regulatory processes. *Journal of Applied Developmental Psychology*, 11(2), 179–193.
- West, M. R. (2012). *Is retaining students in the early grades self-defeating? (CCF brief # 49)*. The Brookings Institution, Center on Children and Families. Retrieved from website <https://www.brookings.edu/wp-content/uploads/2016/06/16-student-retention-west.pdf>.
- Westphal, A., Becker, M., Vock, M., Maaz, K., Neumann, M., & McElvany, N. (2016). The link between teacher-assigned grades and classroom socioeconomic composition: The role of classroom behavior, motivation, and teacher characteristics. *Contemporary Educational Psychology*, 46, 218–227. <https://doi.org/10.1016/j.cedpsych.2016.06.004>.
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R., & Davis-Kean, P. (2006). Development of achievement motivation. In W. Damon, & N. Eisenberg (Eds.). *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (pp. 933–1002). (6th ed.). New York, NY: John Wiley.
- Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, 39(1), 1–37. <https://doi.org/10.1111/j.1745-3984.2002.tb01133.x>.
- Witmer, S. M., Hoffman, L. M., & Nottis, K. E. (2004). Elementary teachers' beliefs and knowledge about grade retention: How do we know what they know? *Education*, 125, 173–194.
- Wong, Y.-L., & Zhou, Y. (2017). The operation of grade retention and educational inequality: Who gets retained in Macao and why? *Asia Pacific Journal of Education*, 37, 150–162. <https://doi.org/10.1080/02188791.2016.1216828>.
- Wu, W., West, S. G., & Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of Educational Psychology*, 102, 135–152. <https://doi.org/10.1037/a0016664>.
- Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, 12(4), 695–719. <https://doi.org/10.1177/1094428108327450>.
- Zimmermann, S., Gehrler, K., Artelt, C., & Weinert, S. (2012). *The assessment of reading speed in grade 5 and grade 9*. Bamberg, DE: University of Bamberg.