



Unraveling the Relationship Between Teacher-Assigned Grades, Student Personality, and Standardized Test Scores

Andrea Westphal*, Miriam Vock and Julia Kretschmann

Department of Education, University of Potsdam, Potsdam, Germany

OPEN ACCESS

Edited by:

Cheng Yong Tan,
The University of Hong Kong,
Hong Kong

Reviewed by:

Jesus Alfonso Daep Datu,
The Education University
of Hong Kong, Hong Kong
Qianqian Pan,
The University of Hong Kong,
Hong Kong

*Correspondence:

Andrea Westphal
andrea.westphal@uni-potsdam.de

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 09 November 2020

Accepted: 08 February 2021

Published: 19 March 2021

Citation:

Westphal A, Vock M and
Kretschmann J (2021) Unraveling
the Relationship Between
Teacher-Assigned Grades, Student
Personality, and Standardized Test
Scores. *Front. Psychol.* 12:627440.
doi: 10.3389/fpsyg.2021.627440

The Big Five personality traits play a major role in student achievement. As such, there is consistent evidence that students that are more conscientious receive better teacher-assigned grades in secondary school. However, research often does not support the claim that students that are more conscientious similarly achieve higher scores in domain-specific standardized achievement tests. Based on the Invest-and-Accrue Model, we argue that conscientiousness explains to some extent why certain students receive better grades despite similar academic accomplishments (i.e., achieving similar scores in domain-specific standardized achievement tests). Therefore, the present study examines to what extent the relationship between student personality and teacher-assigned grades consists of direct as opposed to indirect associations (via subject-specific standardized test scores). We used a representative sample of 14,710 ninth-grade students to estimate these direct and indirect pathways in mathematics and German. Structural equation models showed that test scores explained between 8 and 11% of the variance in teacher-assigned grades in mathematics and German. The Big Five personality traits in students additionally explained between 8 and 10% of the variance in grades. Finally, the personality-grade relationship consisted of direct ($0.02 \mid \beta \mid \leq 0.27$) and indirect associations via test scores ($0.01 \mid \beta \mid \leq 0.07$). Conscientiousness explained discrepancies between teacher-assigned grades and students' scores in domain-specific standardized tests to a greater extent than any of the other Big Five personality traits. Our findings suggest that students that are more conscientious may invest more effort to accomplish classroom goals, but fall short of mastery.

Keywords: Big Five, student personality, teacher-assigned grades, grading practice, conscientiousness, mathematics, German, secondary school

INTRODUCTION

Student personality, especially conscientiousness, is crucial for academic achievement in secondary school (Andersen et al., 2020; Westphal et al., 2020b). Personality research has documented the necessity of examining the relationships between student personality and different indicators of student achievement (Oswald et al., 2004). While six-factor models (e.g., HEXACO, Ashton et al., 2004) and three-factor models

(e.g., Paunonen et al., 2003) of personality emerged in cross-cultural research in particular, the literature on educational research has focused on the popular Big Five personality traits. As such, a number of studies has examined whether the Big Five personality traits in students differentially relate to different achievement outcomes, such as standardized test scores and teacher-assigned grades (Spengler et al., 2013; Tetzner et al., 2019). Overall, the most consistent relationships emerged between secondary-school students' conscientiousness and their teacher-assigned grades (Poropat, 2009). Theoretical models, such as the Invest-and-Accrue Model (Hill and Jackson, 2016) "consider[s] the potential for similar mechanisms and pathways to link conscientiousness to life outcomes across multiple domains" (p. 141). However, while students' conscientiousness is linked to their teacher-assigned grades across multiple domains, students' conscientiousness is less consistently associated with scores in standardized achievement tests (e.g., Meyer et al., 2019).

Research on teachers' grading practice has meanwhile established that a substantial amount of variance in teacher-assigned grades can be accounted for by students' standardized test scores (25 to 35%; Bowers, 2011). Moreover, discrepancies between these two indicators of student achievement (i.e., teacher-assigned grades and standardized test scores) can be explained by students' work habits (Hochweber et al., 2014; Westphal et al., 2016). In the light of these results, similar mechanisms that link conscientiousness to important life outcomes (Hill and Jackson, 2016), like high academic engagement, may be differentially relevant for different indicators of achievement. The present study therefore aims to combine both strands of research, i.e., research on teachers' grading practice and on personality research. As such, we examine the amount of variance in teacher-assigned grades that can be explained by students' standardized test scores, using a large and representative sample of ninth-grade students from the German National Educational Panel Study (NEPS). More importantly, we probe whether the Big Five personality traits in students, especially students' conscientiousness, significantly help explain discrepancies between teacher-assigned grades and students' domain-specific standardized test scores and thus why some students get more favorable grades than other students with similar test scores.

Personality and Academic Achievement

Personality traits describe an individual's typical patterns of cognitions, affects, and behaviors as they occur across various situations and contexts (Fleeson, 2001). The five-factor model—known as the "Big Five"—is the most established model for representing personality differences and consists of five broad domains: Openness, conscientiousness, extraversion, agreeableness, and emotional stability (McCrae and Costa, 1997; for six-factor models, e.g., HEXACO, see Ashton et al., 2004; for three-factor models, see e.g., Paunonen et al., 2003). These five dimensions have been replicated for different age groups including children and adolescents (Soto et al., 2008). These Big Five personality traits in students are highly relevant for academic achievement in secondary school and have been linked to students' scores

in standardized achievement tests (Spengler et al., 2013; Andersen et al., 2020), teacher-assigned grades (Laidra et al., 2007), and grade retention (Westphal et al., 2020b). Empirical studies identified different Big Five personality traits as important precursors for academic achievement in terms of standardized test scores (openness: e.g., Spengler et al., 2013; Waiyavutti, 2019; conscientiousness: e.g., Andersen et al., 2020; extraversion and agreeableness: e.g., Tetzner et al., 2019; emotional stability: e.g., Spengler et al., 2013) and teacher-assigned grades (openness: e.g., Steinmayr and Spinath, 2007; conscientiousness: e.g., Laidra et al., 2007; Spinath et al., 2010; Spengler et al., 2013; Rahafar et al., 2016; Zhang and Ziegler, 2016; Waiyavutti, 2019; Westphal et al., 2020a; extraversion: e.g., Steinmayr and Spinath, 2007; Spinath et al., 2010; agreeableness: e.g., Steinmayr and Spinath, 2007; emotional stability: e.g., Spinath et al., 2010). Yet, there is some evidence from meta-analytical studies that students' conscientiousness may be of particular relevance for teacher-assigned grades in secondary school (Poropat, 2009). In contrast, empirical results are more heterogeneous and the picture is less clear for standardized test scores.

Teacher-Assigned Grades and Standardized Test Scores as Indicators of Different Achievement Aspects

The question of the extent to which teacher-assigned grades actually measure student achievement has stimulated educational research for many years (Borghans et al., 2016). Professional guides encourage teachers to rely primarily on students' achievement of the relevant learning goals in class when assigning final report card grades (e.g., Brookhart, 2004; Linn and Miller, 2005; Brookhart et al., 2016). Thus, Brookhart (2004) underlines that, first and foremost, grades should function as a reliable signifier of student achievement in class for the students themselves and their parents. McMillan (2001) surveyed more than 1,400 secondary-school teachers who reported that they take, to some extent, students' work habits, effort, and participation in class into account when assigning grades. Empirical research comparing teacher-assigned grades to students' scores in large-scale standardized tests, which measured domain-specific achievement (e.g., Bowers, 2011; Hochweber et al., 2014; Westphal et al., 2016), indicates that students' scores in standardized achievement tests explain 25 to 35% of the variance in teacher-assigned grades (Bowers, 2011; although some studies report lower associations, e.g., Westphal et al., 2020a). Willingham et al. (2002) showed that discrepancies between teacher-assigned grades and standardized test scores can be attributed to the observable behavior of students, such as participation in class and homework completion. Other studies similarly found that student effort or motivation explain, to a substantial degree, why students achieve better or worse teacher-assigned grades than their scores in standardized tests would suggest (Hochweber et al., 2014; Westphal et al., 2016). Thus, teacher-assigned grades and standardized test scores measure slightly different aspects of achievement and achievement-related behavior, respectively.

The Process by Which Students' Personality Traits Might Affect Teacher-Assigned Grades and Test Scores

Several potential pathways linking personality to student achievement have been described (Poropat, 2014). Conscientiousness has been linked to effortful control (De Pauw et al., 2009) and self-control (MacCann et al., 2009), both of which could result in more time spent learning (Poropat, 2014). This is in line with the Invest-and-Accrue Model stating that highly conscientious individuals would be “devoting current resources (energy, time, assets) with the intent of achieving current or future success in an important life domain” (Hill and Jackson, 2016, p. 142). Experiencing that their striving indeed pays off will—in the sense of a feedback loop—strengthen their conscientiousness even more. Gaining benefits will subsequently enable them to invest energy in multiple domains. Beyond conscientiousness, openness is also seen as an investment trait, but in the sense that it is characterized by “stable individual differences in the tendency to seek out, engage in, enjoy, and continuously pursue opportunities for effortful cognitive activity” (von Stumm et al., 2011, p. 225). Thus, more open students may typically mobilize more cognitive effort which is in line with findings showing that openness may improve achievement in school because open students may be better able to utilize learning strategies (Blickle, 1996), may be better critical thinkers (Bidjerano and Dai, 2007), and may have higher achievement motivation (Tempelaar et al., 2007). More emotionally stable students, it has been suggested, may be less easily distracted from learning tasks, leading in turn to more time spent on these tasks and subsequently higher levels of achievement (De Raad and Schouwenburg, 1996).

The Invest-and-Accrue Model could be utilized to explain why students' conscientiousness is consistently associated with teacher-assigned grades, but less consistently so with standardized test scores. As such, students that are more conscientious may invest more resources to accomplish classroom goals than less conscientious students. As a consequence, students who achieve similar scores in standardized achievement tests, but are more conscientious, may receive better grades. This is in line with the notion that teachers reward the classroom behavior of highly conscientious students by assigning those students grades that are more positive than their actual academic achievement might warrant (e.g., Spengler et al., 2013). Accordingly, De Raad and Schouwenburg (1996) highlighted that different facets of conscientiousness—i.e., attributes such as self-controlled, well-organized, and mature—have been cited as being core characteristics in studies describing profiles of what might be considered an ideal student. The question arises, then, of whether students' with specific Big Five personality traits achieve better grades because of teachers' grading practices or because of the relation of these characteristics to their actual achievement. Spengler et al. (2013) raised a similar question, using two large samples of secondary school students to examine whether students' personality traits showed similar associations

with teacher-assigned grades when compared to standardized test scores. Their results showed that teacher-assigned grades were mainly associated with conscientiousness, while standardized test scores were mainly related to openness (Spengler et al., 2013). The authors suggested that their findings could be attributed to the fact that “teachers evaluate not only students' abilities but also their studiousness when making decisions about grades” (Spengler et al., 2013, p. 621).

The Present Study

The present study expands the bounds of research on teachers' grading by unraveling the relationship between teacher-assigned grades, student scores in standardized tests, and students' personality traits. We examined whether discrepancies between teacher-assigned grades and students' standardized achievement scores were systematically attributable to any of the Big Five personality traits in students. We utilized data from the National Educational Panel Study (Blossfeld et al., 2011) to explore this relationship in a representative sample of ninth-grade students in Germany, looking separately for teacher-assigned grades in mathematics and German.

As outlined in the introduction, on the one hand, there is a broad research literature showing that students' scores in standardized achievement tests explain a substantial amount of variance in teacher-assigned grades (e.g., Bowers, 2011). Discrepancies between teacher-assigned grades and scores in domain-specific standardized tests, on the other hand, can be partially attributed to students' work habits (e.g., Willingham et al., 2002; Hochweber et al., 2014). The higher investment in classroom goals that highly conscientious students show—as we would expect based on the Invest-Accrue-Model—could help explain why these students receive better teacher-assigned grades despite similar scores in standardized achievement tests. One method that may be useful in disentangling the role of personality in students' levels of ability and studiousness is the analysis of indirect effects. Given these findings, we formulated the following hypotheses.

Hypothesis 1: We expected that students' scores in standardized achievement tests in mathematics and German would explain a substantial amount of variance in teacher-assigned grades in mathematics and German.

Hypothesis 2: We hypothesized that the Big Five personality traits in students—especially conscientiousness—would explain teacher-assigned grades in mathematics and German over and above students' scores in standardized achievement tests in mathematics and German.

Hypothesis 3: We expected that the association between students' Big Five personality traits and teacher-assigned grades would be at least partially explained by students' scores in standardized achievement tests.

MATERIALS AND METHODS

Sample

We employed data from the National Educational Panel Study (NEPS; Blossfeld et al., 2011), a longitudinal multi-cohort study, administered in all 16 German federal states, and which focused on educational processes and competence development. The NEPS is carried out under the supervision of the German Federal Commissioner for Data Protection and Freedom of Information (BfDI), in coordination with the German Standing Conference of the Ministers of Education and Cultural Affairs (KMK), and the Educational Ministries of the respective federal states. Other published manuscripts have used data from the NEPS to examine teacher-assigned grades (Bayer and Zinn, 2018; Brandt et al., 2020). Bayer and Zinn (2018) have used data from the NEPS to test the relationship between teacher-assigned grades and students' standardized test scores without incorporating students' Big Five personality traits. Brandt et al. (2020) examined the associations between Big Five personality traits and teacher-assigned grades based on the NEPS, but they did not take standardized test scores in reading and mathematics into account when predicting teacher-assigned grades. The data protection unit of the Leibniz Institute for Educational Trajectories (LifBi) has approved all instruments and data collection procedures. Participation in the NEPS was voluntary. The parents of all participants gave informed consent for their children to participate in the study, which they are able to revoke at any time. The sample comprised $N = 16,425$ ninth-grade students assessed in the first and second wave, i.e., at the beginning and at the end of the 2010–11 school year. For our study, we only included students in general secondary schools (but not remedial schools), who were surveyed along with at least five of their classmates (in other words, we excluded classes with class sizes of five or less students). Our final sample therefore consisted of $N = 14,710$ students in 922 classes. Fifty percent of students were female. The mean age of students was 14.7 years ($SD = 0.7$). On average, there were 16 students per class.

Measures

In our analyses, we incorporated teacher-assigned grades, scores in standardized achievement tests in mathematics and German as well as indicators of students' Big Five personality characteristics.

Teacher-Assigned Grades

Students reported their mid-year grades obtained in grade nine in mathematics and German (Wave 2). In the German school system, students receive numeric grades from 1 to 6, where a grade of 1 marks excellent achievement, and a grade of 6 reflects unsatisfactory achievement. For our analyses, we reverse-coded the grades so that positive regression coefficients reflect positive relationships. Thus, a grade of 6 denoted excellent achievement, while a grade of 1 reflected unsatisfactory achievement.

Scores in Standardized Achievement Tests

Standardized test scores for mathematics (Wave 1) and reading (Wave 2) were available. To obtain standardized test scores for mathematics, the NEPS applied 22 items

from four different content areas (quantity; space and shape; change and relationships; data and chance) and six different cognitive processes (arguing; communicating; modeling; problem-solving; representing; applying technical skills) with content areas distributed across the items (Durchhardt and Gerdes, 2013). The items had a multiple-choice response format (and a short constructed response in the case of one item). Although different content areas were assessed, a unidimensional partial credit model was found to fit the data well (Durchhardt and Gerdes, 2013). WLE reliability was high (0.79; Durchhardt and Gerdes, 2013).

The NEPS measured students' reading competence using 33 items that assessed three cognitive requirements (finding information in text; drawing text-related conclusions; reflecting and assessing) and different text functions (informational texts; instructional texts; advertising texts; commentaries or arguments; literary texts; Haberkorn et al., 2012). A partial credit model fitted the data well and WLE reliability was high (0.75; Haberkorn et al., 2012).

Personality Characteristics

Students' personality traits were measured using the German 10-item version (BFI-10; Rammstedt and John, 2007) of the Big Five Inventory (John et al., 1991) in Wave 1. The BFI-10 assesses each of the Big Five traits using two items (one positively poled and one negatively poled) from the original scales of the Big Five Inventory. It determines to what extent 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). Reasonable stabilities for all BFI-10 subscales have been demonstrated (test-retest reliabilities; Gosling et al., 2003). Each item was answered on a 5-point scale (1 = does not apply at all to 5 = fully applies).

Statistical Analyses

We used Structural Equation Models (SEMs) with MPlus 7.4 (Muthén and Muthén, 2015) to test our hypotheses. We first specified an exploratory structural equation modeling approach (ESEM), which combines confirmatory (CFA) and exploratory factor analysis (EFA) in one model (Marsh et al., 2014). To account for well-known difficulties in replicating the five-factor structure we modeled students' Big Five personality traits as EFA factors. To identify all factors, their variance was fixed to one and all loadings were estimated freely. To account for "yes-saying" (acquiescence), we specified a response style factor (Aichholzer, 2014). We used oblique geomin rotation. Each item showed the highest loading on the corresponding factor. Thus, the factor solution reflected the theoretically assumed five-factor structure. We have outlined the factor loadings and factor correlations in **Supplementary Table 1**. Beyond the latent Big Five dimensions, we included manifest indicators to capture teacher-assigned grades and standardized test scores. To estimate our model, we used maximum likelihood estimation. To handle the dependency that resulted from the nested data structure

(with students clustered within classes), we used the option `type = complex` in Mplus (adjusting the standard errors for the model parameters; classroom was the level 2 unit). We used bootstrapping to estimate confidence intervals. To estimate how much variance in teacher-assigned grades can be explained by standardized test scores, we specified a model that included grades in mathematics and German as dependent variables and students' standardized test scores in mathematics and German as independent variables (Model 1). We modeled correlations between our two dependent variables as well as correlations between the different standardized test scores. Next, we took the personality traits into account to test whether students' Big Five personality traits would explain incremental variance in teacher-assigned grades over and above standardized test scores (Model 2). When specifying Model 2, we also estimated indirect effects of each Big Five personality trait on teacher-assigned grades via standardized test scores (using the command `model indirect` in Mplus). Model fit was evaluated based on well-established criteria (CFI > 0.95–0.97; RMSEA < 0.05–0.08, SRMR < 0.05–0.10; Hu and Bentler, 1998; Schermelleh-Engel et al., 2003). The average percentage of missing values was 2.51%. We used the robust maximum likelihood (MLR) estimator in conjunction with the full-information maximum-likelihood approach (FIML; Enders, 2001) to obtain appropriate estimates and standard errors.

RESULTS

Descriptive statistics between all study variables are presented in **Table 1**. In addition, bivariate correlations between the latent Big Five personality traits in students and teacher-assigned grades and standardized test scores are depicted in **Figure 1**. Here we can see that conscientiousness was most closely related to teacher-assigned grades in both subjects ($r = 0.19$ for mathematics to $r = 0.21$ for German). For openness and teacher-assigned grades in German, we found a correlation coefficient of $r = 0.18$,

while all other correlations between teacher-assigned grades and personality traits were $|r| = 0.12$ or below. In addition, reading test scores showed the highest correlations with openness ($r = 0.19$), agreeableness ($r = -0.13$), and conscientiousness ($r = -0.10$). Standardized math test scores were most closely associated with emotional stability ($r = 0.17$), agreeableness ($r = -0.16$), and conscientiousness ($r = -0.13$). All other correlations between students' standardized test scores and personality traits did not exceed $|r| = 0.04$.

Variance in Teacher-Assigned Grades Explained by Students' Standardized Test Scores

The first aim of this study was to estimate to what extent students' standardized test scores would explain the variance in teacher-assigned grades. We therefore specified a SEM in which teacher-assigned grades in mathematics and German were regressed on standardized test scores (**Table 2**, Model 1). The model showed a satisfactory fit, CFI = 0.986, RMSEA = 0.036, SRMR = 0.018. We found that students' standardized test scores in mathematics predicted 11.9% of the variance in teacher-assigned grades in mathematics. In German, standardized scores in reading predicted 8.7% of the variance in teacher-assigned grades.

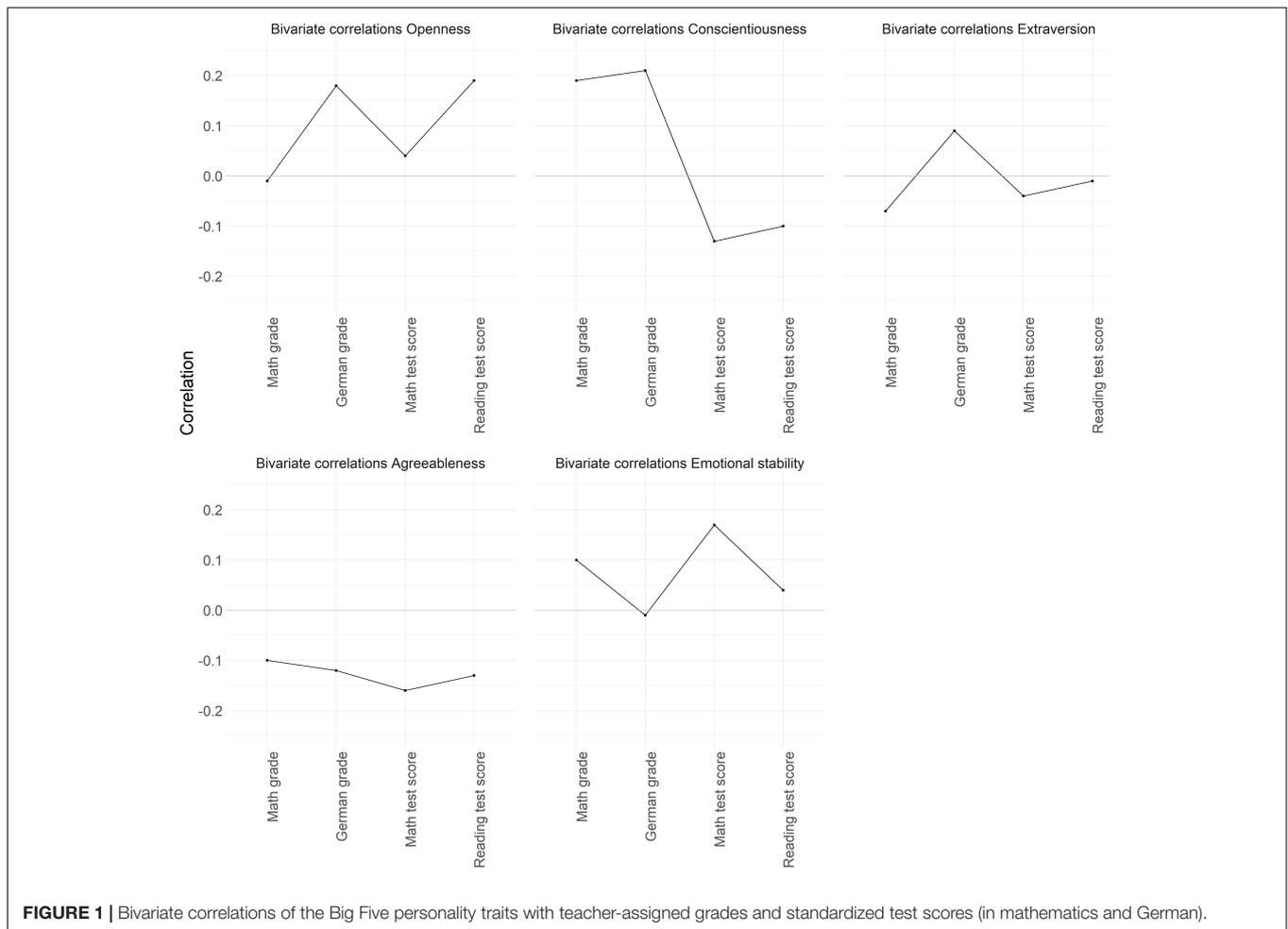
Incremental Variance in Teacher-Assigned Grades Explained by Student Personality

In a second step, we additionally took the Big Five personality traits in students into account (**Table 2**, Model 2). The model fit was acceptable, CFI = 0.979, RMSEA = 0.034, SRMR = 0.015. Students' personality traits additionally explained 7.7% of the variance in teacher-assigned grades in mathematics and 9.7% of the variance in teacher-assigned grades in German. When controlling for Big Five personality traits, the effect of students' standardized test scores on teacher-assigned grades remained stable in both subjects ($\beta = 0.34$ for mathematics test scores;

TABLE 1 | Descriptive statistics for BFI-10 items, teacher-assigned grades and standardized test scores.

	M	SD	Range	Skewness/ Kurtosis	1	2	3	4	5	6	7	8	9	10	11	12	13
1 O Item 1	3.73	1.08	1–5	−0.49/−0.42													
2 O Item 2*	2.78	1.75	1–5	0.21/−1.09	−0.28												
3 C Item 1	3.54	0.87	1–5	−0.19/−0.36	0.08	−0.08											
4 C Item 2*	3.23	1.38	1–5	−0.16/−0.80	0.04	0.11	−0.37										
5 E Item 1	3.51	0.90	1–5	−0.22/−0.31	0.17	−0.01	0.05	0.00									
6 E Item 2*	2.63	1.24	1–5	0.13/−0.69	−0.06	0.03	0.06	0.03	−0.46								
7 A Item 1	3.38	1.07	1–5	−0.32/−0.37	0.08	−0.04	0.07	0.00	0.12	0.03							
8 A Item 2*	2.85	1.07	1–5	0.17/−0.50	0.05	0.08	−0.10	0.22	0.14	−0.14	−0.11						
9 ES Item 1	3.29	1.14	1–5	−0.15/−0.57	0.07	0.06	0.04	0.14	0.14	−0.03	0.04	0.00					
10 ES Item 2*	2.84	1.19	1–5	0.16/−0.65	−0.01	0.00	0.01	0.04	−0.22	0.31	0.08	−0.01	−0.26				
11 G math	4.02	1.05	1–6	−0.11/−0.39	−0.01	0.01	0.18	−0.09	−0.03	0.07	0.00	0.00	0.02	−0.05			
12 G German	4.13	0.70	1–6	−0.09/−0.03	0.09	−0.10	0.18	−0.12	0.09	−0.04	−0.01	0.04	−0.04	−0.04	0.39		
13 T math	0.02	1.48	−4.37–4.62	0.69/0.78	0.06	0.02	0.00	0.13	0.00	0.06	−0.01	0.10	0.07	−0.08	0.36	0.22	
14 T reading	−0.02	1.60	−4.75–3.30	0.13/0.05	0.14	0.04	−0.01	0.09	0.04	0.04	−0.03	0.08	−0.01	−0.03	0.19	0.29	0.54

O, openness; C, conscientiousness; E, extraversion; A, agreeableness; ES, emotional stability; G, grade; T, standardized test score. *Inverse item (not reverse-coded).



$\beta = 0.27$ for reading test scores). In addition, we found that students' conscientiousness was the personality trait that best predicted teacher-assigned grades in both subjects ($\beta = 0.27$ in mathematics; $\beta = 0.26$ in German). In mathematics, all other personality traits also statistically significantly predicted teacher-assigned grades to a significant degree ($\beta = -0.04$ for openness; $\beta = -0.07$ for extraversion; $\beta = -0.12$ for agreeableness; $\beta = 0.08$ for emotional stability). Thus, students with a higher conscientiousness and emotional stability obtained better grades, while higher openness, extraversion and agreeableness were associated with lower grades. In German, openness ($\beta = 0.11$), extraversion ($\beta = 0.06$), and agreeableness ($\beta = -0.19$) were also statistically significant predictors of teacher-assigned grades. Higher openness, higher extraversion, and lower agreeableness were associated with better grades in the subject.

Analyses of Indirect Effects of the Personality-Grade Relationships via Standardized Test Scores

To examine whether the association between students' personality traits and teacher-assigned grades might be partially explained by students' standardized test scores, we

modeled indirect effects of personality traits on grades via test scores when specifying our SEM (Table 2, Model 2). All indirect effects were statistically significant, but standardized regression coefficients were small ($|\beta| \leq 0.07$). We found negative indirect effects of agreeableness ($\beta = -0.07/-0.04$ in mathematics/German), conscientiousness ($\beta = -0.03$ in mathematics and German), and extraversion ($\beta = -0.03/-0.01$ in mathematics/German) and positive indirect effects of openness ($\beta = 0.03/0.06$ in mathematics/German) and emotional stability ($\beta = 0.06/0.02$ in mathematics/German) on teacher-assigned grades in mathematics and German. These results remained virtually identical when taking gender, socioeconomic status, and first-language status as covariates into account (see Supplementary Table 2).

DISCUSSION

Empirical research points to the importance of students' Big Five personality traits in educational outcomes (e.g., Poropat, 2014). In secondary school, students' conscientiousness in particular seems to be highly relevant, with teachers assigning more favorable grades to students that are more conscientious

TABLE 2 | Predicting teacher-assigned grades in mathematics and German.

	Mathematics			German		
	β	p	95% CI	β	p	95% CI
Model 1						
Test score ^a	0.35	0.000	[0.33, 0.36]	0.30	0.000	[0.27, 0.32]
R ²		11.9%			8.7%	
CFI			0.986			
RMSEA			0.036			
SRMR			0.018			
Model 2						
Direct effects						
Test score ^a	0.34	0.000	[0.32, 0.36]	0.27	0.000	[0.24, 0.29]
Openness	-0.04	0.003	[-0.07, -0.01]	0.11	0.000	[0.07, 0.14]
Conscientiousness	0.27	0.000	[0.24, 0.30]	0.26	0.000	[0.23, 0.28]
Extraversion	-0.07	0.000	[-0.10, -0.05]	0.06	0.000	[0.04, 0.09]
Agreeableness	-0.12	0.000	[-0.16, -0.08]	-0.19	0.000	[-0.25, -0.13]
Emotional stability	0.08	0.000	[0.05, 0.11]	0.02	0.143	[-0.01, 0.06]
Indirect effects						
Openness	0.03	0.000	[0.02, 0.05]	0.06	0.000	[0.05, 0.07]
Conscientiousness	-0.03	0.000	[-0.04, -0.02]	-0.03	0.000	[-0.04, -0.02]
Extraversion	-0.03	0.000	[-0.04, -0.02]	-0.01	0.001	[-0.02, -0.01]
Agreeableness	-0.07	0.000	[-0.09, -0.05]	-0.04	0.000	[-0.06, -0.03]
Emotional stability	0.06	0.000	[0.05, 0.08]	0.02	0.001	[0.01, 0.03]
R ²		19.6%			18.4%	
CFI			0.979			
RMSEA			0.034			
SRMR			0.015			

Coefficients are standardized. ^aMathematics/reading test score.

(e.g., Spengler et al., 2013). In contrast, empirical results on the role of secondary-school students' conscientiousness for the scores students receive in domain-specific standardized tests are mixed, with a large number of studies indicating that conscientiousness is of no particular relevance to these test scores (e.g., Meyer et al., 2019). The Invest-and-Accrue Model argues that similar principles explain why conscientiousness affords benefits in multiple outcomes—i.e., conscientious individuals investing more energy, which leads to success, which in turn encourages conscientious behavior (Hill and Jackson, 2016). However, these principles may differentially affect different indicators of educational success. As such, more diligent work habits seem to explain why some students receive better grades despite similar scores in domain-specific standardized-achievement tests (e.g., Willingham et al., 2002). The present study aims to clarify whether students' Big Five personality traits, especially conscientiousness, explain discrepancies in teacher-assigned grades and domain-specific standardized test scores.

Firstly, we found that standardized test scores explained between 8.7 and 11.9% of the variance in teacher-assigned grades, which is less than what has been found for other countries, for instance the United States (25 to 35%, Bowers, 2011). Previous research conducted in Germany similarly found that standardized test scores explain a smaller amount of variance in grades (e.g., 11.6% in a study measuring German secondary school students' achievement in mathematics, Hochweber et al., 2014) than in other countries. In addition, standardized tests

designed to capture the components from the German national educational standards for German and mathematics in secondary school explained between 20 and 28% of the variance in teacher-assigned grades in mathematics and German (Nachtigall, 2015, 2018). The standardized tests that we used for the present study correspond to the German national educational standards for mathematics and German and to the framework in the Program for International Student Assessment (PISA; Institute for Educational Quality Improvement, n.d.; Neumann et al., 2013). Thus, differences in the curricular validity of the standardized tests may be responsible for the fact that standardized test scores in our study explained a smaller amount of variance in grades than in other studies conducted in Germany (Nachtigall, 2015, 2018).

Secondly, we found that the Big Five personality traits in students explained a substantial amount of incremental variance in teacher-assigned grades beyond domain-specific standardized test scores (8% in mathematics and 10% in German). Of the Big Five personality traits, students' conscientiousness and agreeableness were the best predictors of teacher-assigned grades. This finding is in line with the extensive body of literature showing that conscientiousness is particularly important for achieving good grades in secondary school (for meta-analytical evidence, see Poropat, 2009). Based on the Invest-Accrue-Model, this finding may reflect the fact that "conscientious individuals will perform actions that tend to maximize success" in important life domains, such as education (Hill and Jackson, 2016, p. 143). In addition, our finding that agreeableness is

negatively linked to teacher-assigned grades could be explained by research showing that individuals that score high on agreeableness find competitive situations less rewarding, more challenging, and more problematic than individuals low on agreeableness (Graziano et al., 1997). On the other hand, there is an indication that expressive independence—a behavior that puts the focus on expressing one's own ideas, instead of focusing on social responsibility and adjusting to external requirements—is associated with better teacher-assigned grades (Stephens et al., 2014). A possible explanation may therefore be that teachers perceive a more agreeable student as less confident in their own ideas than a less agreeable student with a similar level of achievement, and that teachers then interpret the student's behavior as reflecting lower achievement. The fact that agreeableness is more closely associated with teacher-assigned grades in German than in mathematics may also help explain why findings are more heterogeneous in terms of the role of agreeableness in grading.

Thirdly, our findings indicate that the relationship between the Big Five personality traits in students and teacher-assigned grades is composed of both direct pathways and indirect pathways via standardized test scores. On the one hand, our results therefore support the rationale that students with certain personality traits achieve better scores in a more objective measure of achievement, such as standardized tests, and in turn receive better grades. Our findings (i.e., indirect association of openness and grades) are in line with the notion that students with a higher openness achieve higher scores in achievement tests and receive more favorable grades because they use learning strategies more appropriately, exhibit a higher achievement emotion, and may be better critical thinkers (Blickle, 1996; Bidjerano and Dai, 2007; Tempelaar et al., 2007). In addition, our findings are consistent with research by Graziano et al. (1997) showing that participants high on agreeableness may perceive competition as threatening, which could result in lower achievement and, as a consequence, lower grades. Our results (i.e., indirect association of emotional stability and grades) may also support the idea that more emotionally stable students are less vulnerable to being distracted and therefore spend more time on tasks (De Raad and Schouwenburg, 1996), which may in turn lead to better test scores and higher grades.

Overall, especially when considering conscientiousness, the strength of the direct association between students' conscientiousness and teacher-assigned grades ($\beta \geq 0.26$) was much larger than the strength of the indirect association via test scores ($|\beta| = 0.03$). Our results hold when we additionally take students' gender, socioeconomic background, and first-language status into account (**Supplementary Table 2**). Thus, this finding backs up the argument that—based on the Invest-and-Accrue Model (Hill and Jackson, 2016)—highly conscientious students are more invested in classroom goals. Such an investment may not pay off in higher domain-specific achievement as would be evident in standardized achievement tests. However, it may contribute to the well-established finding that students that are more conscientious receive better grades (e.g., Laidra et al., 2007; Spinath et al., 2010; Spengler et al., 2013). In addition, our results are consistent with and add to research on teachers'

grading practices, which has repeatedly shown that observable student behavior—which is indicative of students' effort and motivation, like classroom participation and homework completion—explains discrepancies between students' grades and test scores (Willingham et al., 2002; see also Hochweber et al., 2014; Westphal et al., 2016). These findings could also be explained by teachers' grading strategies being partially responsible for the relationship between teacher-assigned grades and students' personality traits, such as conscientiousness. This may indicate that teachers value the classroom behavior of highly conscientious students more highly, which they then reward by assigning those students grades that are more positive than their actual academic achievement might warrant (for a similar suggestion see Spengler et al., 2013). Teachers themselves also reported that they took these behaviors into account favorably when assigning grades (McMillan, 2001). Indeed, in informal conversations, teachers described their higher-performing students as “caring more about doing well in school,” having “the will’ to complete their homework, pay[ing] attention in class, and otherwise work[ing] hard” (Duckworth et al., 2015, p. 14). According to Poropat, the social desirability of some Big Five personality traits “may also be important, resulting in a positive halo effect that may raise teachers' ratings of children's academic performance” (Poropat, 2014, p. 242). To sum up, students' conscientiousness explains, to some extent, the discrepancies between teacher-assigned grades and students' scores in domain-specific standardized tests.

Limitations and Future Research

The present study has several limitations. For a start, we cannot rule out the possibility that the direct links found in our study between students' personalities and the grades that teachers assign them might be partially attributable to higher student achievement in subdomains that were not measured by standardized achievement tests. Thus, the use of broad standardized tests adapted even more specifically to the given school's curriculum would be a welcome addition to future research.

Secondly, we captured student personality using a short scale. Short scales sometimes exhibit problems in terms of their factorial validity. The use of short scales, which assess personality traits more narrowly, might also lead researchers to underestimate the role of personality in educational outcomes, including academic achievement (Credé et al., 2012). At the same time, Rammstedt and John (2007) found that their short form gauged 70% of the variance of the long version of the Big Five Inventory and, among different short scales, the short version of the Big Five Inventory by Rammstedt and John (2007) demonstrated the highest criterion validity (Credé et al., 2012). While our results are in line with previous studies that use long versions of personality scales (Poropat, 2014), future research should incorporate longer versions when testing the role of personality in teachers' grading, given the potential limitations of short personality scales. In addition, future studies should shed light on the question of how different subdimensions

of Big Five personality traits relate to different indicators of academic achievement.

Thirdly, our sample consisted of secondary school students in the ninth grade and we cannot generalize our findings to other age groups. There is some indication that different personality traits play a role for teacher-assigned grades in primary than in secondary school (Laidra et al., 2007) and therefore more research with primary students and other age groups is needed. To add to this, the majority of studies have been conducted in Western countries (for exceptions see e.g., Rahafar et al., 2016; Zhang and Ziegler, 2016; Waiyavutti, 2019). Future studies should therefore focus on student personality and its role for achievement in non-Western societies. As the Big Five personality traits are not always replicable in samples that are less well-educated, democratic, rich, industrialized, and non-Western (e.g., Gurven et al., 2013), researchers should complement recent findings by applying three-factor and six-factor models of personality (e.g., Paunonen et al., 2003; Ashton et al., 2004). Finally, the social relations model (e.g., Kenny et al., 2006) suggests that the relationship between the teacher (i.e., the judge) and the student (i.e., the target being judged) is a relevant aspect for interpersonal phenomena and, thus, it might be promising if future studies on teachers' grading practice include additional measures that gauge the relationship between teacher and student.

Practical Implications and Conclusion

In this study, we were able to disentangle the roles of the Big Five personality traits in students and standardized test scores in the grades that teachers assign them. Our findings suggest that conscientiousness explains discrepancies between teacher-assigned grades and scores in domain-specific standardized tests to a greater extent than any of the other Big Five personality traits. Although, according to the Invest-and-Accrue Model (Hill and Jackson, 2016), similar principles may link conscientiousness to important educational outcomes, these pathways may differentially affect teacher-assigned grades and standardized test scores. Our results also correspond to the idea that “grades are used to reward students for exerting effort to learn material, even if students fall short of mastery” (Kelly, 2008, p. 32). In terms of theory-development, the present study underlines the advantages of integrating empirical findings from two strands of research (personality research and research into teachers' grading practices) when deriving conclusions about the very nature and development of student achievement in order to better understand “what's in a grade” (Bowers, 2011, p. 141).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: The Research Data Center at the Leibniz Institute for Educational Trajectories (RDC-LifBi) prepares and disseminates the survey data of the German National Educational Panel Study in the form of

Scientific Use Files to the scientific community. The datasets are available for download from the study website: <https://www.neps-data.de/Data-Center/Data-Access/Download>.

ETHICS STATEMENT

The NEPS study is conducted under the supervision of the German Federal Commissioner for Data Protection and Freedom of Information (BfDI) and in coordination with the German Standing Conference of the Ministers of Education and Cultural Affairs (KMK) and – in the case of surveys at schools – the Educational Ministries of the respective Federal States. All data collection procedures, instruments and documents were checked by the data protection unit of the Leibniz Institute for Educational Trajectories (LifBi). The necessary steps are taken to protect participants' confidentiality according to national and international regulations of data security. Participation in the NEPS study is voluntary and based on the informed consent of participants. This consent to participate in the NEPS study can be revoked at any time.

AUTHOR CONTRIBUTIONS

AW was responsible for the planning, data analysis, and writing of this manuscript, and primary investigator on the DFG funded project. JK and MV contributed through intellectual collaboration and manuscript revisions. All authors contributed to the article and approved the submitted version.

FUNDING

This 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.

ACKNOWLEDGMENTS

We thank Ben Fergusson for his editorial assistance. This study used data from the National Educational Panel Study (NEPS)-Starting Cohort Grade 9 (doi: 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 of organizations.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.627440/full#supplementary-material>

REFERENCES

- Aichholzer, J. (2014). Random intercept EFA of personality scales. *J. Res. Pers.* 53, 1–4. doi: 10.1016/j.jrjp.2014.07.001
- Andersen, S. C., Gensowski, M., Ludeke, S. G., and John, O. P. (2020). A stable relationship between personality and academic performance from childhood through adolescence. An original study and replication in hundred-thousand-person samples. *J. Pers.* 88, 925–939. doi: 10.1111/jopy.12538
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., et al. (2004). A six-factor structure of personality-descriptive adjectives. Solutions from psycholexical studies in seven languages. *J. Pers. Soc. Psychol.* 86, 356–366. doi: 10.1037/0022-3514.86.2.356
- Bayer, M., and Zinn, S. (2018). *Zur Entwicklung des Zusammenhangs und der Erklärbarkeit von Leistungen und Kompetenzen bei Schülerinnen und Schülern der Sekundarstufe I (LifBi Working Paper No. 77) [On the Development of the Relationship between Achievement and Competence of Students in Secondary School]*. Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.
- Bidjerano, T., and Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learn. Individ. Differ.* 17, 69–81. doi: 10.1016/j.lindif.2007.02.001
- Blickle, G. (1996). Personality traits, learning strategies, and performance. *Eur. J. Pers.* 10, 337–352.
- Blossfeld, H.-P., Roßbach, H.-G., and von Maurice, J. (eds) (2011). “The national educational panel study: need, main features, and research potential,” in *Education as a Lifelong Process. The German National Educational Panel Study (NEPS) (Zeitschrift für Erziehungswissenschaft, Special Issue 14)*, eds H.-P. Blossfeld, H.-G. Roßbach, and J. von Maurice (Wiesbaden: VS Verlag für Sozialwissenschaften), 5–17.
- Borghans, L., Golsteyn, B. H. H., Heckman, J. J., and Humphries, J. E. (2016). What grades and achievement tests measure. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13354–13359. doi: 10.1073/pnas.1601135113
- Bowers, A. J. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. *Educ. Res. Eval.* 17, 141–159. doi: 10.1080/13803611.2011.597112
- Brandt, N. D., Lechner, C. M., Tetzner, J., and Rammstedt, B. (2020). Personality, cognitive ability, and academic performance: differential associations across school subjects and school tracks. *J. Pers.* 88, 249–265. doi: 10.1111/jopy.12482
- Brookhart, S. M. (2004). *Grading*. Upper Saddle River, NJ: Pearson.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., et al. (2016). A century of grading research: meaning and value in the most common educational measure. *Rev. Educ. Res.* 86, 803–848. doi: 10.3102/0034654316672069
- Credé, M., Harms, P., Niehorster, S., and Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the big five personality traits. *J. Pers. Soc. Psychol.* 102, 874–888. doi: 10.1037/a0027403
- De Pauw, S. S. W., Mervielde, I., and Van Leeuwen, K. G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *J. Abnorm. Child Psychol.* 37, 309–325. doi: 10.1007/s10802-008-9290-0
- De Raad, B., and Schouwenburg, H. C. (1996). Personality in learning and education: a review. *Eur. J. Pers.* 10, 303–336.
- Duckworth, A. L., Shulman, E. P., Mastrorarde, A. J., Patrick, S. D., Zhang, J., Druckman, J., et al. (2015). Will not want: self-control rather than motivation explains the female advantage in report card grades. *Learn. Individ. Differ.* 39, 13–23. doi: 10.1016/j.lindif.2015.02.006
- Durchhardt, C., and Gerdes, A. (2013). *NEPS Technical Report for Mathematics—Scaling Results of Starting Cohort 4 in Ninth Grade*. NEPS Working Paper No. 22. Bamberg: University of Bamberg, National Educational Panel Study.
- Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychol. Methods* 6, 352–370. doi: 10.1037/1082-989X.6.4.352
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: traits as density distributions of states. *J. Pers. Soc. Psychol.* 80, 1011–1027. doi: 10.1037/0022-3514.80.6.1011
- Gosling, S. D., Rentfrow, P. J., and Swann, W. B. Jr. (2003). A very brief measure of the Big Five personality domains. *J. Res. Pers.* 37, 504–528. doi: 10.1016/S0092-6566(03)00046-1
- Graziano, W. G., Hair, E. C., and Finch, J. F. (1997). Competitiveness mediates the link between personality and group performance. *J. Pers. Soc. Psychol.* 73, 1394–1408. doi: 10.1037/0022-3514.73.6.1394
- Guven, M., Von Rueden, C., Massenkoff, M., Kaplan, H., and Lero Vie, M. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager-farmers in the Bolivian Amazon. *J. Pers. Soc. Psychol.* 104:354. doi: 10.1037/a0030841
- Haberkorn, K., Pohl, S., Hardt, K., and Wiegand, E. (2012). *NEPS Technical Report for Reading—Scaling Results of Starting Cohort 4 in Ninth Grade*. NEPS Working Paper No. 16. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Hill, P. L., and Jackson, J. J. (2016). The invest-and-accrue model of conscientiousness. *Rev. Gen. Psychol.* 20, 141–154. doi: 10.1037/gpr0000065
- Hochweber, J., Hosenfeld, I., and Klieme, E. (2014). Classroom composition, classroom management, and the relationship between student attributes and grades. *J. Educ. Psychol.* 106, 289–300. doi: 10.1037/a0033829
- Hu, L.-T., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol. Methods* 3, 424–453. doi: 10.1037/1082-989X.3.4.424
- Institute for Educational Quality Improvement (n.d.). *Kompetenzstufenmodelle [Proficiency Level Models]*. Available online at: <https://www.iqb.hu-berlin.de/bista/ksm> (accessed July 30, 2019).
- John, O. P., Donahue, E. M., and Kentle, R. L. (1991). *The Big Five Inventory: Versions 4a and 5a*. Berkeley, CA: University of California, Institute of Personality and Social Research.
- Kelly, S. (2008). What types of students' effort are rewarded with high marks? *Sociol. Educ.* 81, 32–52. doi: 10.1177/003804070808100102
- Kenny, D. A., Kashy, D. A., and Cook, W. L. (2006). *Dyadic Data Analysis*. New York, NY: Guilford Press.
- Laidra, K., Pullmann, H., and Allik, J. (2007). Personality and intelligence as predictors of academic achievement: a cross-sectional study from elementary to secondary school. *Pers. Individ. Differ.* 42, 441–451. doi: 10.1016/j.paid.2006.08.001
- Linn, R., and Miller, M. (2005). *Measurement and Assessment in Teaching*. Upper Saddle River, NJ: Pearson Prentice Hall.
- MacCann, C., Duckworth, A. L., and Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learn. Individ. Differ.* 19, 451–458. doi: 10.1016/j.lindif.2009.03.007
- Marsh, H. W., Morin, A. J. S., Parker, P. D., and Kaur, G. (2014). Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annu. Rev. Clin. Psychol.* 10, 85–110. doi: 10.1146/annurev-clinpsy-032813-153700
- McCrae, R. R., and Costa, P. T. (1997). Personality trait structure as a human universal. *Am. Psychol.* 52, 509–516.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educ. Meas.* 20, 20–32. doi: 10.1111/j.1745-3992.2001.tb00055.x
- Meyer, J., Fleckenstein, J., Retelsdorf, J., and Köller, O. (2019). The relationship of personality traits and different measures of domain-specific achievement in upper secondary education. *Learn. Individ. Differ.* 69, 45–59. doi: 10.1016/j.lindif.2018.11.005
- Muthén, L. K., and Muthén, B. O. (2015). *Mplus Version 7.4 Software [Internet]*. Los Angeles: Statmodel.
- Nachtigall, C. (2015). *Landesbericht. Thüringer Kompetenztests 2015 [State Report. Written Comparison Tests in Thuringia]*. Available online at: <https://www.kompetenztest.de/downloads/kompetenztests/archive> (accessed October 26, 2020).
- Nachtigall, C. (2018). *Landesbericht. Thüringer Kompetenztests 2018 [State Report. Written Comparison Tests in Thuringia]*. Available online at: <https://www.kompetenztest.de/downloads/kompetenztests/archive> (accessed October 26, 2020).
- Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E., and Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *J. Educ. Res.* 5, 80–109.
- Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., and Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *J. Appl. Psychol.* 89, 187–207.
- Paunonen, S. V., Haddock, G., Forsterling, F., and Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *Eur. J. Pers.* 17, 413–433. doi: 10.1002/per.496

- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychol. Bull.* 135, 322–338. doi: 10.1037/a0014996
- Poropat, A. E. (2014). A meta-analysis of adult-rated child personality and academic performance in primary education. *Br. J. Educ. Psychol.* 84, 239–252. doi: 10.1111/bjep.12019
- Rahafar, A., Maghsudloo, M., Farhangnia, S., Vollmer, C., and Randler, C. (2016). The role of chronotype, gender, test anxiety, and conscientiousness in academic achievement of high school students. *Chronobiol. Intern.* 33, 1–9. doi: 10.3109/07420528.2015.1107084
- Rammstedt, B., and 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. *J. Res. Pers.* 41, 203–212. doi: 10.1016/j.jrp.2006.02.001
- Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res.* 8, 23–74.
- Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. (2008). The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *J. Pers. Soc. Psychol.* 94, 718–737. doi: 10.1037/0022-3514.94.4.718
- Spengler, M., Lütke, O., Martin, R., and Brunner, M. (2013). Personality is related to educational outcomes in late adolescence: evidence from two large-scale achievement studies. *J. Res. Pers.* 47, 613–625. doi: 10.1016/j.jrp.2013.05.008
- Spinath, B., Freudenthaler, H. H., and Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Pers. Individ. Differ.* 48, 481–486. doi: 10.1016/j.paid.2009.11.028
- Steinmayr, R., and Spinath, B. (2007). Predicting school achievement from motivation and personality. *Zeitschrift Pädagogische Psychol.* 21, 207–216. doi: 10.1024/1010-0652.21.3.207
- Stephens, N. M., Markus, H. R., and Phillips, L. T. (2014). Social class culture cycles: how three gateway contexts shape selves and fuel inequality. *Annu. Rev. Psychol.* 65, 611–634. doi: 10.1146/annurev-psych-010213-115143
- Tempelaar, D. T., Gijssels, W. H., van der Loeff, S. S., and Nijhuis, J. F. H. (2007). A structural equation model analyzing the relationship of student achievement motivations and personality factors in a range of academic subject-matter areas. *Contemp. Educ. Psychol.* 32, 105–131. doi: 10.1016/j.cedpsych.2006.10.004
- Tetzner, J., Becker, M., and Brandt, N. D. (2019). Personality-achievement associations in adolescence—examining associations across grade levels and learning environments. *J. Pers.* 88, 356–372. doi: 10.1111/jopy.12495
- von Stumm, S., Chamorro-Premuzic, T., and Ackerman, P. L. (2011). “Re-visiting intelligence-personality associations: vindicating intellectual investment,” in *The Wiley-Blackwell Handbook of Individual Differences*, eds T. Chamorro-Premuzic, S. von Stumm, and A. Furnham (Chichester: Wiley-Blackwell), 217–241.
- Waiyavutti, C. (2019). *Development and Validation of Thai Versions of Big Five Measures from the International Personality Item Pool*. Iowa City, IA: University of Iowa, doi: 10.17077/etd.i8ax-txq8
- Westphal, A., Becker, M., Vock, M., Maaz, K., Neumann, M., McElvany, N., et al. (2016). The link between teacher-assigned grades and classroom socioeconomic composition: the role of classroom behavior, motivation, and teacher characteristics. *Contemp. Educ. Psychol.* 46, 218–227. doi: 10.1016/j.cedpsych.2016.06.004
- Westphal, A., Lazarides, R., and Vock, M. (2020a). Are some students graded more appropriately than others? Student characteristics as moderators of the relationships between teacher-assigned grades and test scores in mathematics. *Br. J. Educ. Psychol.* doi: 10.1111/bjep.12397 [Epub ahead of print].
- Westphal, A., Vock, M., and Lazarides, R. (2020b). 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. *J. Appl. Dev. Psychol.* 66:101088. doi: 10.1016/j.appdev.2019.101088
- Willingham, W. W., Pollack, J. M., and Lewis, C. (2002). Grades and test scores: accounting for observed differences. *J. Educ. Measur.* 39, 1–37. doi: 10.1111/j.1745-3984.2002.tb01133.x
- Zhang, J., and Ziegler, M. (2016). How do the big five influence scholastic performance? A big five-narrow traits model or a double mediation model. *Learn. Individ. Differ.* 50, 93–102. doi: 10.1016/j.lindif.2016.07.011

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Westphal, Vock and Kretschmann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.