Received: 6 November 2017 | Revised: 19 May 2018 | Accepted: 31 May 2018

DOI: 10.1111/jopy.12410

ORIGINAL ARTICLE

WILEY

The effects of getting a new teacher on the consistency of personality

Sven Rieger[1] | **Richard Göllner**[1] | **Marion Spengler**[1] | **Ulrich Trautwein**[1] | **Benjamin Nagengast**[1] | **Jeffrey R. Harring**[2] | **Brent W. Roberts**[1,3]

[1]Hector Research Institute of Education Sciences and Psychology, University of Tübingen

[2]Human Development and Quantitative Methodology, University of Maryland

[3]Department of Psychology, University of Illinois at Urbana–Champaign

Correspondence
Sven Rieger, Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastr. 6, 72072 Tübingen, Germany.
Email: sven.rieger@uni-tuebingen.de

Funding Information
LEAD Graduate School & Research Network [GSC1028], funded by Excellence Initiative of the German federal and state governments; European Social Fund and Ministry of Science, Research and the Arts of Baden-Württemberg grant to Marion Spengler; Ministry of Science, Research and the Arts of Baden-Württemberg grant (Az: 33-7532.20/735) to Ulrich Trautwein; The Graduate School of the University of Maryland, College Park and the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63); TRAIN study initiated/funded by grants from Ministries of Education, Youth, and Sports in Baden-Württemberg and Saxony, Germany.

Abstract

Objective: In the present research, we examined the effect of getting a new teacher on consistency in students' personality measures, including trait and social cognitive constructs.

Method: To test the effect of this kind of situational transition, we analyzed two large longitudinal samples ($N = 5{,}628$; $N = 2{,}458$) with quasi-experimental study designs. We used two consistency measures (i.e., rank-order clations and changes in variance over time) to compare students who got a new teacher with students who kept the same teacher.

Results: Multiple-group latent variable analyses showed no differences in the rank-order correlations for the math-related social cognitive constructs of interest, effort, self-concept, self-regulation, anxiety, and the Big Five personality traits. Significantly lower rank-order correlations were found for some of the German- and English-related social cognitive constructs (i.e., effort measures) for the group of students who got a new teacher. Regarding the changes in variance (over time), we found no systematic differences between groups in both studies.

Conclusions: We found partial support for the idea that social cognitive variables are more susceptible to environmental changes (i.e., getting a new teacher) than the Big Five personality traits are.

KEYWORDS
Big Five, consistency, new teacher, social cognitive constructs

1 | INTRODUCTION

It is widely assumed that teachers influence and shape the life experiences of their students (Brophy, 1986; Pianta, 1999; Wentzel, 2002). Students interact with their teachers every day, and, thus, it is not surprising that, for instance, positive student–teacher relationships (e.g., characterized by supportiveness, encouragement of thinking, empathy) are associated with positive school outcomes such as enhanced performance and positive socioemotional development (Cornelius-White, 2007; Hamre & Pianta, 2001; McCormick & O'Connor, 2015). Given the impact of teachers on the development of

young students and the fact that students get new teachers every year in most school systems, an important question follows: What happens to students when they get a new teacher? Students do not know much about their new teacher's expectations and attitudes, and, consequently, such a transition entails a degree of uncertainty that should putatively cause at least a temporary change in students' thoughts, feelings, or behaviors.

In the present investigation, we were interested in the extent to which a variety of personality constructs would be affected when students got a new teacher (i.e., when they underwent a change in situation). Constructs such as interests, self-concepts, and conscientiousness are both interesting and relevant because of the evidence that these types of school-related constructs are important for school performance and school functioning (Eccles & Wigfield, 2002; Kautz, Heckman, Diris, Ter Weel, & Borghans, 2014; Poropat, 2009). To optimally foster them, it is important to understand the nature and malleability of these variables (Bailey, Duncan, Odgers, & Yu, 2017). An indication of malleability is whether and to what extent a new situation (e.g., getting a new teacher) affects these constructs.

Thus, the present article was designed to address three questions related to the effect of getting a new teacher on broadly construed measures of students' psychological functioning (including traits and social cognitive constructs). First, we explored whether the rank-order correlations of students' personality constructs would be affected when students got a new teacher (entered a new situational context). Second, we investigated whether getting a new teacher or not getting a new teacher would be associated with a change in variance in personality constructs over time (i.e., individual differences becoming more or less pronounced). Third, we aimed to determine whether social cognitive variables would be more strongly influenced than the Big Five personality traits when students got a new teacher. For these purposes, we compared the two indices of consistency (i.e., rank-order correlations and changes in variance components over time) between students who had the same teacher and students who got a new teacher for several social cognitive variables (e.g., interests, effort, and self-concepts) and the Big Five personality traits in two large longitudinal studies.

1.1 | A different context: Getting a new teacher

The extent to which different levels of functioning in personality are influenced by environmental factors is a question that has concerned psychologists for decades (Asendorpf & Van Aken, 2003; Funder & Colvin, 1991; Johnson, 1999; Roberts & Pomerantz, 2004). In recent years, research on the composition and meaning of situations has experienced a revival (Rauthmann et al., 2014; Rauthmann, Sherman, & Funder, 2015; Reis, 2008). Similarly, it has been argued that to get a complete picture of personality development, it is essential to understand the influence, or lack thereof, of different contexts and the experiences people have in these contexts (Roberts & Nickel, 2017). This seems especially important for the periods of late childhood and early adolescence because these are the stages that are defined by fundamental changes in youths' lives (Soto & Tackett, 2015), and personality underlies rapid development (Roberts & Pomerantz, 2004).

In this study, we were particularly interested in whether experiencing a new context that entails some uncertainty would have an impact on students' psychological functioning. Many school systems assign students to new teachers each year, and students are faced with such situations many times during their school years. When students get a new teacher, they have to adapt to new situations and circumstances. They have to deal with different attitudes and different teaching styles, and they must establish a relationship with their new teacher. Because students do not know what to expect from a new teacher, such a transition entails some uncertainty. What can be expected to happen to students' outcomes when they experience such a transition? The uncertainty associated with such a transition may result in some instability in the development of students' outcomes due to adaptation processes (e.g., students will change their behavior to meet new expectations). By contrast, having the same teacher for multiple school years could have stabilizing effects (e.g., Sherman, Nave, & Funder, 2010) because students would not have to adjust to a new teacher. Unfortunately, studies have yet to investigate the relation between this common change in students' environments and their psychological functioning, thus making it difficult to make concrete predictions about the effects.

1.2 | What effect does getting a new teacher have on student characteristics?

Personality can be construed at many different levels. For instance, constructs emerging out of a social cognitive framework are by definition malleable and contextualized (Bandura, 2012; Eccles & Wigfield, 2002). Social cognitive constructs (e.g., interests and self-concepts) are understood to be narrow, relevant to specific contexts, and derived almost exclusively from experience and interactions with others (Eccles & Wigfield, 2002; Suls & Mullen, 1982). By contrast, trait-like (e.g., the Big Five personality traits) variables are often assumed to be biologically based, stable, and not amenable to change (Eysenck, 1970; McCrae & Costa, 2013). Moreover, they are often defined as consistent across situations and relatively independent of the context (see, e.g., Funder & Colvin, 1991; Johnson, 1999). The distinction between traits and social cognitive constructs is supported by different theoretical models that conceptualize traits as core characteristics or basic

tendencies and social cognitive variables as surface characteristics or even characteristic adaptations (Asendorpf & Van Aken, 2003; McAdams & Pals, 2006; McCrae & Costa, 2008).

However, a recent review of the conceptual and empirical basis of the distinction between these two classes of constructs found that the division between trait-like variables (e.g., the Big Five) and social cognitive constructs (e.g., values, self-related schemata) was conceptually larger than the empirical data could justify (Kandler, Zimmermann, & McAdams, 2014). In addition, it is uncommon to find both kinds of variables included in the same study (Roberts, 2009), and this makes it even more difficult to adequately compare these constructs. When tracked over equivalent periods of time using the same methods (i.e., self-reports), social cognitive and personality trait constructs were found to show comparable levels of continuity and change in young childhood and early adolescence (Rieger et al., 2017). Such a finding supports the idea that investigating and testing the proposed characteristics (e.g., consistency across time and contexts) behind these factors is essential for understanding the nature of these variables. In the present article, we focused on the assumption of consistency across contexts and tested the extent to which several social cognitive variables and the Big Five personality traits would respond differently to the same type of environmental experience (i.e., getting a new teacher).

1.3 | Characterizing consistency

To quantify consistency in individual differences (over time), researchers typically consider multiple indicators (see, e.g., Caspi, Roberts, & Shiner, 2005; Fleeson & Noftle, 2008; Mõttus, Soto, & Slobodskaya, 2017). Two types of consistency seem most relevant for registering the perturbation in psychological functioning when the environment changes: rank-order correlations and changes in variances over time (Baltes & Nesselroade, 1973; Roberts & Mroczek, 2008). Rank-order correlations refer to the relative placement of individuals from one time point to the next within a group. Changes (increases or decreases) in variances over time reflect whether individual differences in constructs become more or less pronounced across time (Mõttus et al., 2017). Students who transition to a new teacher must adapt to the teacher's expectations and behaviors. However, student–teacher relationships are complex, and both within and across classes, all students will not respond in the same way to a new teacher (see, e.g., person–environment fit; Eccles et al., 1993). A change in teacher, however, would affect the relative placement of students within the group with a new teacher (i.e., some students will increase, whereas other students will decrease within the same class). The relative placement of the students within the group that did not get a new teacher should be unaltered. Consequently, students with a new teacher should show lower rank-order correlations than students who do not get a new teacher.

Also, it would make sense to expect differences in changes in variances over time (i.e., individual differences that become more or less pronounced in either group) between the groups. This expectation emerges both from conceptual models and recent empirical evidence on variance patterns of personality in children and adolescents. Specifically, theoretical models of the ontogenetic factors that contribute to personality development imply or specify that experiences could result in increased variance. In particular, Fraley and Roberts (2005) showed that random life experiences were fundamental to lower levels of stability over time and potentially increased variability. In more sophisticated models of cross-species personality development (Stamps & Biro, 2016), it was found that one of the explicit consequences of random life experiences or perturbations would be a fanning effect on personality over time. Finally, a recent empirical test of the idea that corresponsive relations between personality and life experience would lead to increasing variance in personality over time found evidence for the effect in childhood and adolescence (Mõttus et al., 2017). In particular, Mõttus et al. (2017) found that variance in personality increased from early childhood until early adolescence and that this pattern held for most personality traits from the Big Five. In a follow-up, genetically informed study, it was found that the increasing variance was most likely attributable to unique Gene × Gene or Gene × Environment interactions. Interestingly, but appropriate to the current study, the increase in variance that may result from experience appears to apply only to children and adolescents and not older populations (Mõttus, Allik, Hřebíčková, Kööts-Ausmees, & Realo, 2016). Thus, it is reasonable that the adaptation processes of the students who got a new teacher might lead to larger variation in personality constructs. On the other side, students who did not get a new teacher would already be familiar with their teacher, and, thus, individual differences should be more stable across time. Both processes would result in a difference in changes in variances over time between the groups. Finally, an examination of mean-level change was less relevant in the current analyses because the reaction to a new teacher could be a positive or negative experience, thus leading to a lack of meaningful patterns in mean-level change over time.

1.4 | The present research

In the present research, we examined the effect of getting a new teacher on consistency in students' psychological functioning in two independent, large longitudinal German studies. The German school system provided the opportunity to test the effect of getting a new teacher in an optimal setting. In Germany, teachers typically change classes every 2 years, whereas the composition of these classes remains the same over several years. Thus, by using quasi-experimental designs, we compared two indices of consistency (i.e., 1-year rank-order correlations and changes in variance components

over time) between students who got a new teacher and those who did not on a variety of student characteristics. In Study 1, we analyzed several math-related social cognitive variables (i.e., interest, academic effort, self-regulation, and anxiety). In Study 2, we applied the same approach as in Study 1 but with two noteworthy extensions. First, the social cognitive constructs of interest and effort were complemented by self-concept, and, in addition, we examined potential differences between the school subjects of math, German, and English. Second, we also contrasted the social cognitive constructs with the Big Five personality traits and asked whether the social cognitive variables were more strongly influenced by getting a new teacher than the Big Five personality traits were. On the basis of the often proposed characteristics of both types of variables, social cognitive variables (strong context-sensitivity) should be more influenced by this transitional situation than the Big Five personality traits (which are often believed to be stable across situations and contexts).

2 | STUDY 1

2.1 | Method

2.1.1 | Sample

For Study 1, we analyzed data from the German national extension to the 2003 cycle of the Organisation for Economic Co-operation and Development's Programme for International Student Assessment (PISA; OECD, 2004). In this national extension, a subsample of 15-year-old PISA students and their teachers from Grade 9 took part in an additional assessment in Grade 10. Participation in the study was voluntary. Data collection took part at the ends of both school years.

The complete data set contained $N = 6{,}020$ students in 275 classes. However, to avoid contamination effects due to different class compositions, we had to exclude $n = 392$ students who were not in the same classes at Time 2. This resulted in a data set with $N = 5{,}628$ students (56.4% female) in 259 classes with $n = 1{,}132$ students (20.1%) who had different teachers and $n = 4{,}496$ students (79.9%) who had the same teacher at both time points. The mean age of the students was 15.1 years in Grade 9. In Germany, a "tripartite" system consisting of lower-track schools (*Hauptschule*), intermediate-track schools (*Realschule*), and academic-track schools (*Gymnasium*) is the most common system. However, some federal states also offer multitrack schools (including *Integrierte Gesamtschule* or *Schule mit mehreren Bildungsgängen*), which include several school tracks within the same school. In this study, the students were spread across different school tracks as follows: multitrack schools, $n = 983$ (17.5%); intermediate-track schools, $n = 2{,}199$ (39.1%); and academic-track schools, $n = 2{,}446$ (43.5%). Students from the lower track (*Hauptschule*) were not included because this track ends after Grade 9.

2.1.2 | Instruments

Five items were used to assess individual interest[1] in math (see, e.g., Krapp & Prenzel, 2011). The items targeted interest in and the intrinsic value of math-related activities (Cronbach's alpha: $\alpha_{T1} = .86$, $\alpha_{T2} = .89$). Academic effort in math was also measured with five items (see Pekrun et al., 2005). The items focused on the effort needed to meet subject-specific tasks ($\alpha_{T1} = .78$, $\alpha_{T2} = .80$). Self-regulation in math was measured with five items ($\alpha_{T1} = .70$, $\alpha_{T2} = .75$). The items assessed students' self-regulation of learning goals, use of strategies, and monitoring of learning outcomes (Goetz, 2004). Nine items served to assess math anxiety (Achievement Emotions Questionnaire-Mathematics; Pekrun, Goetz, Titz, & Perry, 2002). Students were instructed to rate how they typically felt when taking tests in math ($\alpha_{T1} = .90$, $\alpha_{T2} = .90$).

2.1.3 | Statistical analyses

We estimated all models in the framework of longitudinal confirmatory factor analyses and used full information maximum likelihood estimation with robust standard errors (MLR; using Mplus Version 7.31; Muthén & Muthén, 1998–2012). Before running the models, we centered the data on the group mean to interpret the correlations and variances as pooled stability and variability within-cluster coefficients. To make the missing at random assumption more plausible, we included several auxiliary variables (Collins, Schafer, & Kam, 2001) in all analyses (i.e., standardized achievement tests, grades). This was done by using the auxiliary option as implemented in Mplus. By using this option, the so-called saturated correlates model (Graham, 2003) can be estimated. In the saturated correlates model, the auxiliary variables are not part of the structural model but are allowed to correlate with the variables. Statistical tests were performed two-sided and used the 5% level of significance.

Measurement invariance

Before addressing our research question, we tested for measurement invariance over time and between groups. If we had failed to establish measurement invariance, we would not have been able to rule out the possibility that differences in correlations over time or between groups were due to the measurement process (Meredith, 1993; Widaman, Ferrer, & Conger, 2010). For this purpose, we specified two models by starting with a model in which weak measurement invariance was imposed over time and between groups (the same factor loadings for each indicator/parcel over time and between groups; called the *liberal model* in the following). In the

next step, we imposed strong measurement invariance over time (the same factor loadings and intercepts over time) and strict measurement invariance between groups (the same factor loadings, intercepts, and residual variances and the same residual correlations between groups; called the *restrictive model* in the following). To evaluate these models, we used fit indices that are commonly used for latent variable models, namely, the Satorra-Bentler-scaled chi-square test (Satorra & Bentler, 2010), the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR; Hu & Bentler, 1998). According to Hu and Bentler (1999), a good fit is indicated by CFI/TLI ≥ 0.95 and RMSEA/SRMR ≤ 0.05. To compare the nested models, we used the chi-square difference test and the recommendations by Chen (2007) and Cheung and Rensvold (2002) as guidelines. The results of their simulation studies suggested that, when testing for invariance in factor loadings, noninvariance is indicated by a change of ≤ 0.01 in the CFI, supplemented by a change of ≤ 0.015 in the RMSEA and ≤ 0.030 in the SRMR. However, in our study, not only did we constrain the factor loadings between groups and over time, but we also constrained multiple parameters within one step (i.e., intercepts over time and intercepts and residual [co]variances between groups).

Comparison of rank-order correlations and variance components

To address the first research question, we compared the rank-order correlations between the groups with the Wald test as implemented in Mplus. To address the second research question, we compared the change (difference) in variance (from T1 to T2) of each group against each other. This was done using the *Z*-test.[2] Before doing this, we applied the natural logarithm transformation of all variances to achieve a (better) approximation of the normality assumption. All coefficients were based on the more restrictive multiple-group latent variable model (see Figure 1). All syntax and outputs can be found at osf.io (https://osf.io/vp2mg).

Parceling and the nested data structure

To reduce model complexity, we decided to build item parcels for each construct. To construct balanced item parcels, we followed the item-to-construct balance parceling strategy (see Little, Cunningham, Shahar, & Widaman, 2002). Thus, we constructed item parcels by considering the relative balance between loadings and intercepts, starting with the highest and then adding the other items to the anchor item in an inverted order. To control for the specific item parcel variance over time, we used the correlated uniqueness approach (Cole, Ciesla, & Steiger, 2007).

In the present investigation, students were nested within classes, resulting in a multilevel structure of the data set. Therefore, students within a class were not independent of each other (i.e., students within classes tend to be more similar than students from different classes; Raudenbush & Bryk, 2002). Not considering this structure could lead to an underestimation of standard errors (see, e.g., Muthén & Satorra, 1995). However, because the intraclass correlations of the variables were rather low (ranging from to .03 to .13), and,

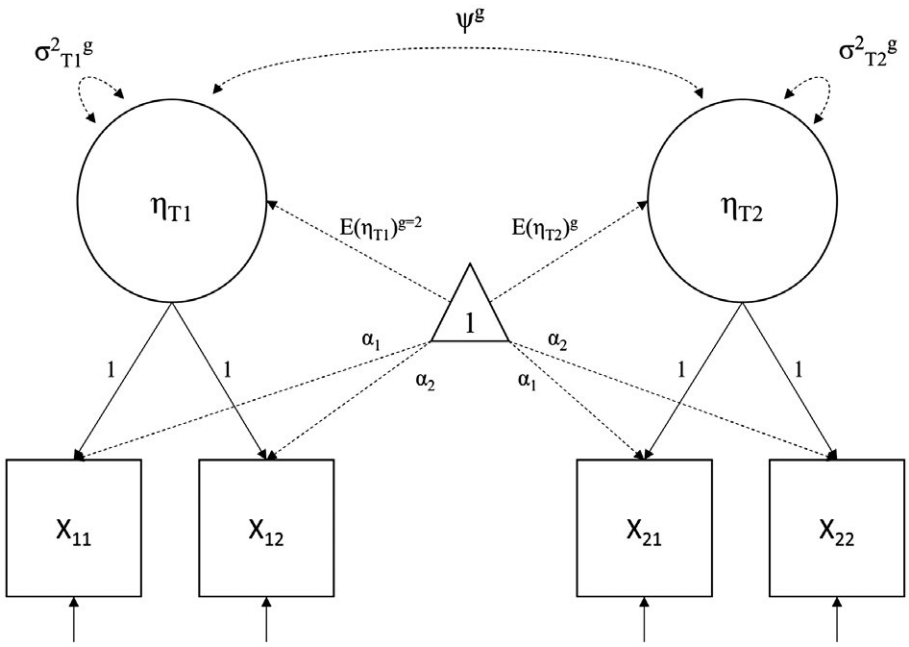

FIGURE 1 Multiple-group latent variable model (restrictive). Error covariances for the same item parcels (correlated uniquenesses) are not depicted. The mean of the first group was fixed to zero so that the model would be identified

more important, because the data were centered on the cluster mean (class mean), there was no need to correct the model parameters for clustering effects (Muthén, 1989; Satorra & Muthén, 1995). Moreover, the standardized model parameters (e.g., correlation coefficients) and standard errors of models based on group-centered data are equivalent to those derived from the within model from an ordinary multilevel model (see also Göllner, Wagner, Eccles, & Trautwein, 2018, p. 7).

2.2 | Results

In the following, we begin by presenting the results of the measurement invariance tests between groups and over time. These are followed by the results of the multiple-group latent variable models. Descriptive results for the overall sample for both time points are displayed in Table S1 in Appendix A1. Furthermore, we did not expect to find any meaningful patterns (e.g., a reaction to a new teacher could be a positive or negative experience) in means over time. In addition, after centering the data on the group mean, there are no mean-level differences left. However, we report all mean-level results from the analyses without group-mean-centering in Table S2 in Appendix A2. In sum, there were no significant differences in mean-level change over time between the groups.

2.2.1 | Establishing measurement invariance

Before addressing our research question, we tested for measurement invariance over time and between groups. For this purpose, we tested a liberal model (i.e., the same factor loadings for each indicator/parcel over time and between groups) against a (more) restrictive model (i.e., the same factor loadings and intercepts over time and the same factor loadings, intercepts, residual variances, and residual correlations between groups). Table S3 in Appendix A3 displays all model fit criteria of all tested models. All models demonstrated a good fit (CFI/TLI > 0.95, RMSEA/SRMR ≤ 0.07) to the data.

With regard to the model comparisons: According to the chi-square difference test, there were no significant differences between the liberal and the more restrictive models on all variables. Moreover, there were no noticeable differences in the fit indices (i.e., CFI, RMSEA, and SRMR) between the two models (except for anxiety: ΔRMSEA = 0.02). Thus, we assumed strong measurement invariance over time and strict measurement invariance between groups.

2.2.2 | Comparison of rank-order correlations

We estimated the 1-year rank-order correlations of the two groups (same teacher [st] vs. new teacher [nt]) and all variance components separately for each construct by means of the (more restrictive) multiple-group latent variable models. The latent rank-order correlations are presented in Table 1.

We found no significant differences in the rank-order correlations for all four constructs: interest ($r_{st} = .75$ vs. $r_{nt} = .74$, $\Delta = -.01$, $T_W = 0.12$, $p = .728$), academic effort ($r_{st} = .58$ vs. $r_{nt} = .56$, $\Delta = -.01$, $T_W = 0.05$, $p = .824$), anxiety ($r_{st} = .72$ vs. $r_{nt} = .74$, $\Delta = .01$, $T_W = 0.23$, $p = .632$), and self-regulation ($r_{st} = .60$ vs. $r_{nt} = .56$, $\Delta = -.04$, $T_W = 0.51$, $p = .474$).

2.2.3 | Comparison of variance components

With regard to the second research question, we compared the change in variance from T1 to T2 between the groups (Δ). All estimated variances as well as the differences between the changes in variances from T1 to T2 are presented in Table 2. In the group that did not get a new teacher (δ_{st}), the variances of math interest and self-regulation increased significantly from T1 to T2 (math interest: $\delta_{st} = 0.16$, $Z = 4.65$, $p < .001$; self-regulation: $\delta_{st} = 0.22$, $Z = 3.28$, $p = .001$). The corresponding changes (δ_{nt}) from T1 to T2 in the group that got a new teacher were not statistically significant (math interest: $\delta_{nt} = 0.12$, $Z = 1.55$, $p = .120$; self-regulation: $\delta_{nt} = 0.10$, $Z = 0.70$, $p = .487$). The differences in the changes in variances between the groups were not statistically significant (math interest: $\Delta = -0.04$, $Z = -0.417$, $p =$

TABLE 1 Results of multiple-group latent variable models (PISA study)

Constructs	Same teacher r_{12}	New teacher r_{12}	Difference in time consistency (Δ)[a]
Social cognitive variables			
Math interest	.75 [.72, .78]	.74 [.68, .79]	–.01 [–.07, .05]
Math effort	.57 [.53, .62]	.56 [.47, .66]	–.01 [–.14, .06]
Math self-regulation	.60 [.54, .66]	.56 [.48, .64]	–.04 [–.08, .01]
Math anxiety	.72 [.70, .75]	.74 [.69, .78]	.01 [–.04, .06]

Note. PISA = Programme for International Student Assessment.
[a]Wald test.

TABLE 2 Results of multiple-group latent variable models (PISA study)

Constructs of δ_{st} and δ_{nt}	Same teacher			New teacher			Difference $(\Delta)^a$ of δ_{st} and δ_{nt}
	Var_{T1}	Var_{T2}	$Var_{T2}–Var_{T1}$ (δ^a)	Var_{T1}	Var_{T2}	$Var_{T2}–Var_{T1}$ (δ^a)	
Social cognitive variables							
Math interest	0.44 [0.41, 0.47]	0.52 [0.49, 0.54]	0.16* [0.09, 0.22]	0.47 [0.41, 0.53]	0.53 [0.47, 0.58]	0.12 [–0.03, 0.27]	–0.04 [–0.20, 0.13]
Math effort	0.24 [0.22, 0.26]	0.26 [0.23, 0.28]	0.06 [–0.06, 0.18]	0.23 [0.19, 0.27]	0.25 [0.2, 0.29]	0.06 [–0.17, 0.30]	0.00 [–0.26, 0.26]
Math self-regulation	0.20 [0.18, 0.22]	0.24 [0.21, 0.27]	0.22* [0.09, 0.35]	0.18 [0.15, 0.22]	0.20 [0.16, 0.25]	0.10 [–0.17, 0.36]	–0.12 [–0.42, 0.17]
Math anxiety	0.43 [0.41, 0.45]	0.44 [0.42, 0.46]	0.02 [–0.03, 0.06]	0.44 [0.4, 0.48]	0.44 [0.4, 0.47]	–0.01 [–0.09, 0.07]	–0.03 [–0.12, 0.07]

Note. PISA = Programme for International Student Assessment. Before subtracting the variances, we applied a natural logarithm transformation.
[a]Z-test. [b]$*p < 0.05$.

.676; self-regulation: $\Delta = –0.12$, $Z = –0.81$, $p = .421$). On all other variables, there were no significant differences in the variances within and between the groups.

2.3 | Discussion

In Study 1, we found no significant differences in the rank-order correlations for the two groups on the four social cognitive constructs (i.e., interest, effort, anxiety, and self-regulation). Thus, we found only a little support for the impact of a new teacher on the rank-order correlations (a change in the relative placement from one time point to the next) of math-related social cognitive variables.

Regarding the variance comparison, we found an increase in the variance in the group that did not get a new teacher for math interest and self-regulation. However, this increase was not statistically significantly different from the change in variance in the group that got a new teacher. Moreover, on all other variables, there were no significant differences in the changes in the variances from T1 to T2 between the groups. Consequently, our results suggest that individual differences in the variables included in this study do not get more pronounced for students who get a new teacher than for students who keep the same teacher.

3 | STUDY 2

3.1 | Method

3.1.1 | Sample

In Study 2, we used data from a large longitudinal German study (Tradition and Innovation in Educational Systems; TRAIN; Jonkmann, Rose, & Trautwein, 2013) that is hosted by the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. TRAIN is a large-scale school achievement study that encompasses four time points (from Grades 5 to 8). The study comprises $N = 3{,}876$ students in 136 classes in 99 schools from two federal states (Baden-Württemberg and Saxony).

To have a design that was comparable to the one used in Study 1 (Study 1 focused on Grades 9 and 10), we used the third and fourth time points from the TRAIN study (involving Grades 7 and 8; called T1 and T2 in the following). In sum, there was complete information for $N = 2{,}458$ students (45.4% male) such that $n = 1{,}546$ (62.9%) had different teachers and $n = 912$ (37.1%) had the same teacher at both time points. In the TRAIN study, the teachers of each class were tracked. Class teachers have different main subjects, and this offered us the opportunity to investigate differential

TABLE 3 Numbers of students and teachers across the different subjects

Main subject taught by teacher	T1	T2	Same teacher at T1 and T2	New teacher at T2	Same teacher with same subject at T1 to T2	New teacher with same subject at T1
Math	832 students (38 teachers)	902 students (40 teachers)	297 students (14 teachers)	535 students (24 teachers)	239 students (11 teachers)	405 students (18 teachers)
German	1,060 students (53 teachers)	881 students (44 teachers)	281 students (13 teachers)	779 students (40 teachers)	259 students (12 teachers)	457 students (24 teachers)
English	279 students (13 teachers)	356 students (17 teachers)	123 students (6 teachers)	156 students (7 teachers)	123 students (6 teachers)	129 students (5 teachers)
Other	287 students (14 teachers)	319 students (17 teachers)	211 students (10 teachers)	76 students (4 teachers)	145 students (7 teachers)	27 students (2 teachers)
Total	2,458 students (118 teachers)	2,458 students (118 teachers)	912 students (43 teachers)	1,546 students (75 teachers)	766 students (36 teachers)	1,018 students (48 teachers)

Note. Solid box represents overall sample size. Dashed boxes represent domain-specific sample sizes.

effects between school subjects.[3] For a detailed overview of the sample size composition (i.e., students per teachers with different main subjects), see Table 3.

The students were spread across the two federal states such that in Baden-Württemberg, $n = 1{,}008$ (36.5%) of the students came from the lower track (*Hauptschule*), and $n = 733$ (26.6%) came from the intermediate track (*Realschule*). The remaining students, $n = 1{,}017$ (36.9%), attended multitrack schools (*Mittelschule*) in Saxony.

3.1.2 | Instruments

Social cognitive variables

The social cognitive constructs (i.e., self-concept, interest, and academic effort) were assessed with four items each in three different school subjects, namely, math, German, and English (see Krapp & Prenzel, 2011; Marsh, 1992; Pekrun et al., 2005; Wigfield & Eccles, 2000). The items were rated on a 4-point Likert scale ranging from 1 (*I do not agree at all*) to 4 (*I agree entirely*). The domain-specific interest items focused on the intrinsic value of and interest in the respective school subject (Cronbach's αs ranged from .68 to .75). The items from the academic effort scales (developed for the TRAIN study) focused on the effort needed to meet subject-specific tasks (αs ranged from .85 to .90). The self-concept items targeted the students' own evaluations of their ability in the respective school subjects (αs ranged from .64 to .86).

Big Five personality traits

The Big Five were measured with the German version (Lang, Lüdtke, & Asendorpf, 2001) of the Big Five Inventory (John, Donahue, & Kentle, 1991). The items were rated on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In line with findings from other studies with the same data set (Rieger et al., 2017; Trautwein et al., 2015), for all Big Five traits, the negatively worded items showed negative or low item-total correlations (all rs $< .22$).[4] Thus, we used only the positively worded items. Cronbach's alpha values ranged from .66 to .80.

3.1.3 | Statistical analyses

We again used multiple-group latent variable models to estimate the 1-year rank-order correlations in both groups (students who had the same teacher vs. students who had different teachers) as well as all variance components (see Figure 1). To compare the differences in the rank-order correlations and the differences in variances (from T1 to T2) between the groups, we again used the Wald and Z-test, respectively (see Casella & Berger, 2002).

The model estimation (full information maximum likelihood estimation with robust standard errors), model fit evaluation criteria, parceling strategy, and the way in which we

dealt with the multilevel structure of the data set were the same as in Study 1.

3.2 | Results

We again begin by presenting the results of the measurement invariance tests between the groups and over time. Then we present the results of the multiple-group latent variable models. Descriptive results for the overall sample for both time points are displayed in Table S4 in Appendix B2. Again, we did not expect to find any meaningful patterns in means over time, and centering the data on the group mean (class mean) erased any mean-level differences that may have been present. We report all mean-level results from the analyses without group-mean-centering in Table S5 in Appendix B3. In sum, for most variables, there were no significant differences in mean-level change over time between the groups. Significant differences emerged for Conscientiousness, Openness, and Extraversion. The differences (Δd) between the groups ranged from |0.03| to |0.09|, thus falling within the range of regular maturation effects in this age range (Conscientiousness: $d = 0.03$; social vitality facet of Extraversion: $d = 0.11$; social dominance facet of Extraversion: $d = 0.20$; Agreeableness: $d = 0.01$; see Roberts, Walton, & Viechtbauer, 2006).

3.2.1 | Establishing measurement invariance

In line with the procedure used in Study 1, we first tested for measurement invariance over time and between groups. Thus, we tested a liberal model (the same factor loadings for each indicator/parcel over time and between groups) against a more restrictive model (the same factor loadings and intercepts over time and the same factor loadings, intercepts, residual variances, and residual correlations between groups). Table S6 in Appendix B4 displays all model fit criteria of all tested models. All models demonstrated a good fit (CFI/TLI > 0.95, RMSEA/SRMR ≤ 0.06) to the data.

When we compared the liberal and restrictive models, the restrictive model fit the data significantly worse for the variables of Neuroticism ($\Delta\chi^2 = 17.61$, $\Delta df = 7$, $\Delta p = .013$) and Agreeableness ($\Delta\chi^2 = 19.53$, $\Delta df = 7$, $\Delta p = .001$). For all remaining variables, there were no significant differences between the models. Differences in the fit indices were found for the variables of interest in German (ΔRMSEA = 0.02), effort in English (ΔRMSEA = 0.04), and Neuroticism (RMSEA = 0.03). However, all fit indices (i.e., CFI, TLI, RMSEA, and SRMR) were still above or below the cut-off values of 0.95 or 0.06, respectively. Consequently, we were able to assume strong measurement invariance over time and strict measurement invariance between groups.

TABLE 4 Results of multiple-group latent variable models (TRAIN study)

Constructs	Same teacher Stability r_{12}	New teacher Stability r_{12}	Difference in time consistency (Δ)[a]
Social cognitive variables			
Math interest	.52 [.36, .68]	.45 [.31, .60]	–.06 [–.27, .15]
German interest	.54 [.37, .70]	.28 [.07, .49]	–.26 [–.52, .01]
English interest	.37 [.06, .68]	.23 [–.06, .52]	–.14 [–.60, .33]
Math effort	.57 [.47, .68]	.55 [.44, .64]	–.03 [–.17, .11]
German effort	.58 [.49, .66]	.37 [.22, .52]	–.21* [–.38, –.04]
English effort	.58 [.44, .71]	.28 [.05, .51]	–.30* [–.56, –.03]
Math self-concept	.63 [.48, .77]	.64 [.56, .73]	.02 [–.15, .19]
German self-concept	.81 [.67, .95]	.68 [.56, .81]	–.12 [–.29, .04]
English self-concept	.69 [.54, .84]	.72 [.54, .90]	.03 [–.19, .25]
Big Five personality traits			
Conscientiousness	.53 [.44, .62]	.50 [.43, .57]	–.03 [–.15, .09]
Neuroticism	.45 [.33, .57]	.47 [.38, .56]	.02 [–.13, .17]
Openness	.52 [.40, .64]	.51 [.44, .59]	–.01 [–.15, .13]
Agreeableness	.52 [.40, .65]	.44 [.33, .56]	–.08 [–.24, .08]
Extraversion	.63 [.49, .76]	.52 [.43, .61]	–.11 [–.26, .05]

Note. TRAIN = Tradition and Innovation in Educational Systems.
[a]Wald test.[b]*$p < .05$.

TABLE 5 Results of multiple-group latent variable models (TRAIN study)

Constructs	Same teacher			New teacher			Difference (Δ) [a] of δ_{st} and δ_{nt}
	Var_{T1}	Var_{T2}	Var_{T2}–Var_{T1} (δ^a)	Var_{T1}	Var_{T2}	Var_{T2}–Var_{T1} (δ^a)	
Social cognitive variables							
Math interest	0.30 [0.23, 0.37]	0.27 [0.17, 0.36]	–0.13 [–0.44, 0.18]	0.26 [0.20, 0.32]	0.36 [0.27, 0.46]	0.33* [0.04, 0.63]	0.46* [0.05, 0.88]
German interest	0.32 [0.24, 0.40]	0.25 [0.14, 0.36]	–0.23 [–0.64, 0.19]	0.29 [0.21, 0.36]	0.22 [0.16, 0.28]	–0.27 [–0.61, 0.07]	–0.04 [–0.57, 0.48]
English interest	0.35 [0.17, 0.53]	0.33 [0.16, 0.50]	–0.05 [–0.61, 0.51]	0.38 [0.24, 0.51]	0.16 [0.11, 0.21]	–0.87* [–1.32, –0.42]	–0.82* [–1.46, –0.18]
Math effort	0.33 [0.20, 0.46]	0.38 [0.25, 0.52]	0.14 [–0.08, 0.36]	0.29 [0.24, 0.34]	0.38 [0.33, 0.44]	0.29 [0.09, 0.50]	0.15 [–0.15, 0.44]
German effort	0.40 [0.28, 0.52]	0.40 [0.23, 0.56]	–0.02 [–0.39, 0.34]	0.38 [0.30, 0.45]	0.29 [0.22, 0.35]	–0.27* [–0.52, –0.02]	–0.25 [–0.68, 0.19]
English effort	0.40 [0.21, 0.60]	0.36 [0.24, 0.48]	–0.13 [–0.65, 0.40]	0.29 [0.26, 0.32]	0.34 [0.27, 0.41]	0.16 [–0.09, 0.41]	0.28 [–0.28, 0.85]
Math self-concept	0.45 [0.36, 0.54]	0.49 [0.39, 0.60]	0.09 [–0.04, 0.22]	0.45 [0.37, 0.54]	0.52 [0.41, 0.64]	0.14 [–0.16, 0.45]	0.06 [–0.26, 0.36]
German self-concept	0.22 [0.12, 0.33]	0.18 [0.12, 0.24]	–0.21 [–0.56, 0.13]	0.19 [0.12, 0.25]	0.15 [0.11, 0.20]	–0.20 [–0.55, 0.16]	0.01 [–0.38, 0.41]
English self-concept	0.43 [0.30 0.56]	0.50 [0.33, 0.66]	0.15 [–0.35, 0.65]	0.31 [0.21, 0.42]	0.42 [0.22, 0.62]	0.29 [–0.14, 0.72]	0.14 [–0.50, 0.79]
Big Five personality traits							
Conscientiousness	0.29 [0.24, 0.34]	0.22 [0.19, 0.25]	–0.28* [–0.42, –0.14]	0.24 [0.21, 0.28]	0.24 [0.21, 0.27]	–0.03 [–0.18, 0.13]	0.26* [0.05, 0.46]
Neuroticism	0.19 [0.15, 0.23]	0.18 [0.15, 0.21]	–0.04 [–0.24, 0.16]	0.22 [0.19, 0.26]	0.25 [0.22, 0.28]	0.13 [–0.04, 0.31]	0.17 [–0.08, 0.43]
Openness	0.24 [0.20, 0.28]	0.21 [0.18, 0.24]	–0.13 [–0.31, 0.05]	0.24 [0.21, 0.27]	0.23 [0.20, 0.26]	–0.03 [–0.19, 0.12]	0.10 [–0.13, 0.33]
Agreeableness	0.22 [0.18, 0.25]	0.19 [0.15, 0.24]	–0.13 [–0.34, 0.08]	0.18 [0.14, 0.21]	0.18 [0.15, 0.21]	0.03 [–0.16, 0.21]	0.15 [–0.11, 0.42]
Extraversion	0.22 [0.18, 0.26]	0.20 [0.17, 0.24]	–0.09 [–0.27, 0.08]	0.22 [0.18, 0.25]	0.22 [0.19, 0.24]	0.00 [–0.16, 0.16]	0.09 [–0.13, 0.32]

Note. TRAIN = Tradition and Innovation in Educational Systems. Before subtracting the variances, we applied a natural logarithm transformation.
[a]*Z*-test.[b]*$p < .05$.

3.2.2 | Comparison of rank-order correlations

We estimated the 1-year rank-order correlations and all variance components of both groups (same teacher [st] vs. new teacher [nt]) separately for each construct by means of the more restrictive multiple-group latent variables. The results of the 1-year rank-order correlations are displayed in Table 4.

For the interest measures, we found slightly lower rank-order correlations for the group of students who got a new teacher in all three subjects. However, these differences were not statistically significant (math: $r_{st} = .52$ vs. $r_{nt} = .45$, $\Delta = -.06$, $T_W = -0.36$, $p = .546$; German: $r_{st} = .54$ vs. $r_{nt} = .28$, $\Delta = -.26$, $T_W = 3.71$, $p = .054$; English: $r_{st} = .37$ vs. $r_{nt} = .23$, $\Delta = -.14$, $T_W = 0.33$, $p = .563$). On the effort measures, students who had a new teacher showed statistically significantly lower rank-order correlations than the students who had the same teacher in German ($r_{st} = .58$ vs. $r_{nt} = .37$, $\Delta = -.21$, $T_W = 6.08$, $p = .014$) and English ($r_{st} = .58$ vs. $r_{nt} = .28$, $\Delta = -.30$, $T_W = 4.77$, $p = .029$), but this finding did not hold in math ($r_{st} = .57$ vs. $r_{nt} = .55$, $\Delta = -.03$, $T_W = 0.17$, $p = .681$). For the self-concept measures, we found no significant differences in the rank-order correlations between the two groups (math: $r_{st} = .63$ vs. $r_{nt} = .65$, $\Delta = .02$, $T_W = 0.05$, $p = .829$; German: $r_{st} = .81$ vs. $r_{nt} = .68$, $\Delta = -.12$, $T_W = 2.17$, $p = .141$; English: $r_{st} = .69$ vs. $r_{nt} = .72$, $\Delta = .03$, $T_W = 0.07$, $p = .787$).

In a final step, we compared the rank-order correlations between the groups on the Big Five personality traits. We did not find any statistically significant differences between the two groups on any of the Big Five traits.

3.2.3 | Comparison of variance components

With regard to our second research question, paralleling Study 1, we compared the change in variance from T1 to T2 between the groups (Δ) separately for each construct. All estimated variances as well as the differences between the variances are reported in Table 5. For the social cognitive variables, we found significant differences in the changes in variances between the groups for math interest and interest in English. In the group that did not get a new teacher (δ_{st}), the variances of math interest and interest in English did not significantly decrease from T1 to T2 (math interest: $\delta_{st} = -0.13$, $Z = -0.84$, $p = .401$; interest in English: $\delta_{st} = -0.05$, $Z = -0.18$, $p = .859$). However, the corresponding changes (δ_{nt}) from T1 to T2 in the group that got a new teacher showed statistically significant increases or decreases (math interest: $\delta_{nt} = 0.33$, $Z = 2.19$, $p = .028$; interest in English: $\delta_{nt} = -0.87$, $Z = 3.76$, $p < .001$). The differences in the changes in variance between the groups were statistically significant (math interest: $\Delta = 0.46$, $Z = 2.20$, $p = .028$; interest in English: $\Delta = -0.82$, $Z = -2.50$, $p = .012$). For the Big Five personality traits, we found a significant difference in the change in variance in Conscientiousness between the groups. In the group that did not get a new teacher (δ_{st}), the variance in Conscientiousness showed a statistically significant decrease from T1 to T2 ($\delta_{st} = -0.28$, $Z = 3.92$, $p < .001$). The corresponding change (δ_{nt}) from T1 to T2 in the group that got a new teacher was not statistically significant ($\delta_{nt} = -0.03$, $Z = -0.32$, $p = .747$). However, the difference in the change in variance between the groups was statistically significant ($\Delta = 0.26$, $Z = 2.46$, $p = .014$). On all other variables, we found no significant differences in the change in the variance from T1 to T2 between the groups. However, two additional issues should be noted. First, the variance in effort in German showed a statistically significant decrease in the group that got a new teacher. However, this change was not statistically significantly different from the change in variance in the other group. Second, the variances of the Big Five traits got descriptively smaller in the group that did not get a new teacher, and no systematic pattern emerged in the group that got a new teacher (see Table 5).

3.3 | Discussion

In line with Study 1, we found no differences in the rank-order correlations for the math-related social cognitive variables between students who got a new teacher and students who did not get a new teacher. However, on the social cognitive constructs that are related to German and English (i.e., effort in English and German), students who got a new teacher showed significantly lower rank-order correlations (i.e., their relative placement within the group changed) than students who did not get a new teacher. On the Big Five personality traits, there were no significant differences in the rank-order correlations between the groups. In sum, our results suggest that social cognitive variables are slightly more influenced by a changing context (i.e., getting a new teacher) than the Big Five personality traits.

Regarding the variance comparison, we found only three (of 14) significant differences in the change in the variances over time between the groups (i.e., math interest, interest in English, and Conscientiousness). Moreover, there was a significant decrease in the variance over time within the group that got a new teacher in effort in German. On the basis of these contradictory patterns, we concluded that there were no systematic patterns in the increases or decreases in variances over time within or between groups.

4 | GENERAL DISCUSSION

In the present investigation, we examined consistency in students' psychological functioning in two independent longitudinal studies of students who got a new teacher in comparison with those who did not get a new teacher. For this purpose, we compared the 1-year rank-order

correlations and variance components of a variety of personality variables between students who got new teachers and those who did not. By analyzing two large data sets, we found (a) no differences in the rank-order correlations of math-related social cognitive constructs between the two groups. Furthermore, we found (b) no significant differences in rank-order correlations for the Big Five personality traits. However, we found (c) significantly lower consistencies across time for the group of students who got a new teacher on some of the German- and English-related social cognitive constructs (i.e., effort measures). Finally, regarding the comparison of variances, we found (d) no systematic differences (increases or decreases) within or between groups across the two studies.

4.1 | Cross-situational consistency in students' characteristics

The extent to which students' personality is influenced by environmental factors is a central question in personality research (Asendorpf & Van Aken, 2003; Funder & Colvin, 1991; Johnson, 1999; Roberts & Pomerantz, 2004). We examined the effect of getting a new teacher—a situation that students face many times in their school careers—on consistency in a variety of student characteristics. We were particularly interested in whether social cognitive constructs would be more susceptible to this kind of situation than trait constructs such as the Big Five. Multiple theoretical models can distinguish between two groups of variables, namely, core traits/characteristics (also called *basic tendencies*) and surface characteristics (also called *characteristic adaptations*; Asendorpf & Van Aken, 2003; McAdams & Pals, 2006; McCrae & Costa, 2008). On the basis of group allocation, the constructs are conceptualized as either stable and consistent across contexts or unstable and contextualized. Past research showed that social cognitive and personality trait constructs are comparable regarding levels of continuity and change in a stable environment (Rieger et al., 2017).

In the present research, we focused on consistency across different environmental experiences (i.e., getting a new teacher) and found partial support for the idea that social cognitive variables are more susceptible to environmental changes than the Big Five personality traits are. The strongest effects were found for the effort measures (in the subjects English and German), a finding that coincides with the strong relation between teacher behavior and student effort (Pianta, Hamre, & Allen, 2012). However, the math-related social cognitive constructs (self-concept, interest, effort, self-regulation, and anxiety) were unaffected by the change in teachers. One explanation could be that math-related constructs are more entwined with a person's ability, and this might lead to a certain robustness against environmental influences such as getting a new teacher.

In line with the theoretical assumptions about core traits/basic tendencies (McCrae & Costa, 2008), we found no effect of getting a new teacher on the consistency of the Big Five personality traits. Thus, it can be concluded that there is no marked impact of this particular different contextual situation (i.e., getting a new teacher) on the consistency of the broad personality traits. This finding is in line with Funder and Colvin's (1991) reasoning. They suggested that cross-situational consistency also depends on the "level" of behavior and that higher-level (global personality) traits show more consistency than middle- or lower-level behaviors (which are more concrete; see Leikas, Lönnqvist, & Verkasalo, 2012). Consequently, focusing on the lower levels of the Big Five personality traits can offer suitable insights for identifying and understanding the processes behind the broader traits. Research has not paid much attention to the extent to which the facets of personality traits are stable and changeable, especially in adolescence. However, initial insights can be derived from the study by Jackson et al. (2009). They found that not all facets of Conscientiousness change in a similar way from early to later adulthood. However, how stable and context-sensitive the facets are has yet to be addressed by research.

Finally, our results do not mean that teachers do not influence the development of students' psychological functioning. The present study tested solely the extent to which a variety of personality constructs would respond (differently) to the same environmental experience (i.e., a new teacher). The beneficial impact of positive teacher characteristics (e.g., supportiveness) on social cognitive variables such as engagement and interest are well documented (Frenzel, Goetz, Pekrun, & Watt, 2010; Pianta et al., 2012; Wentzel, 1998, 2002; Wigfield & Eccles, 1992). The extent to which teacher characteristics (e.g., attitude toward orderliness) influence students' personality development has not been the subject of (much) research until now, but it is definitely worth exploring.

5 | LIMITATIONS AND FUTURE DIRECTIONS

Although both studies used a quasi-experimental design and large samples and examined a variety of social cognitive constructs as well as all Big Five personality traits, some limitations should be kept in mind when interpreting the results. First, both studies relied on global and retrospective self-reports. Roberts (2018), for instance, criticized the assessment of global, retrospective reports of personality constructs. He claimed that this assessment method fails to provide a sufficient measure of the constructs in general. Instead of measuring global, retrospective reports of personality constructs,

Roberts (2018) suggested that only states should be assessed over time (e.g., by tracking information from smartphones, tablets, etc.). With enough assessments of states over time, one could both extract trait-level patterns and track change as it occurs in response to environmental change. However, self-reports are commonly used to measure constructs such as effort, individual interest, and self-concept, and also the Big Five personality traits. Given the focus on a variety of social cognitive constructs, self-reports are perhaps one of the most valid measures for adequately capturing students' feelings and perceptions. Moreover, using the same method helped to maintain the comparability of the two construct classes. Nevertheless, it is important for future studies to examine context sensitivity with different measures as well (e.g., behavioral measures). Second, we investigated pure between-person time consistency indicators. Considering also intrapersonal processes (e.g., within-person time consistency; Hamaker, Kuiper, & Grasman, 2015; Sherman et al., 2010) is essential for understanding the nature of the constructs. Finally, although it was one aim of the present research to compare the social cognitive constructs with the personality traits, these classes differ in their conception and granularity. Whereas the Big Five traits are defined as broad and domain-general constructs, social cognitive constructs are conceptualized as narrow and domain specific. Researchers should consider examining lower-order models (facets) of each of the Big Five domains to achieve a more fine-grained understanding of the Big Five (e.g., the extent to which the facets are stable and contextualized) in future studies.

ACKNOWLEDGMENTS

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this manuscript was supported by several grants. Sven Rieger is a doctoral student at the LEAD Graduate School & Research Network [GSC1028], which is funded by the Excellence Initiative of the German federal and state governments. This research was supported by a grant to Marion Spengler funded by the European Social Fund and the Ministry of Science, Research and the Arts of Baden-Württemberg. This work was supported by the Ulrich Trautwein from the Ministry of Science, Research and the Arts of Baden-Württemberg (Az: 33-7532.20/735). This research project was also supported by the Graduate School of the University of Maryland, College Park and the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63). The TRAIN study was initiated and is funded by grants from the Ministries of Education, Youth, and Sports in Baden-Württemberg and Saxony, Germany.

CONFLICT OF INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

NOTES

[1]There is no standardized and widely accepted instrument for the assessment of individual interest. Multiple questionnaires with good face validity and good psychometric quality have been used (see Renninger & Hidi, 2008).

[2]In Mplus, the standard errors for this test are based on the delta method (see Casella & Berger, 2002).

[3]Owing to the different school subjects, it was not possible to consider more than two time points because the sample sizes of the cells became too small to obtain reliable estimates of the rank-order correlations and especially the variance components.

[4]These results probably had to do with response biases such as acquiescence, midpoint responding, or extreme responding. Using self-reports on the Big Five personality traits in young children is very challenging and is currently under discussion in the literature (see, e.g., Göllner et al., 2017; Soto, John, Gosling, & Potter, 2008).

ORCID


Sven Rieger http://orcid.org/0000-0002-0339-0695

REFERENCES

Asendorpf, J. B., & Van Aken, M. A. G. (2003). Personality-relationship transaction in adolescence: Core versus surface personality characteristics. *Journal of Personality*, *71*, 629–666. https://doi.org/10.1111/1467-6494.7104005

Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*, 7–39. https://doi.org/10.1080/19345747.2016.1232459

Baltes, P. B., & Nesselroade, J. R. (1973). The developmental analysis of individual differences on multiple measures. In J. R. Nesselroade, & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological issues* (pp. 219–251). New York, NY: Academic Press. https://doi.org/10.1016/B978-0-12-515650-9.50017-4

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*, 9–44. https://doi.org/10.1177/0149206311410606

Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, *41*, 1069–1077. https://doi.org/10.1037/0003-066X.41.10.1069

Casella, G., & Berger, R. L. (2002). *Statistical inference*. Pacific Grove, CA: Duxbury/Thomson Learning.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, *12*, 381–398. https://doi.org/10.1037/1082-989X.12.4.381

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*, 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, *77*, 113–143. https://doi.org/10.3102/003465430298563

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, *48*, 90–101. https://doi.org/10.1037/0003-066X.48.2.90

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eysenck, H. J. (1970). *The structure of human personality*. London: Routledge.

Fleeson, W., & Noftle, E. E. (2008). Where does personality have its influence? A supermatrix of consistency concepts. *Journal of Personality*, *76*, 1355–1386. https://doi.org/10.1111/j.1467-6494.2008.00525.x

Fraley, R. C., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review*, *112*, 60–74. https://doi.org/10.1037/0033-295X.112.1.60

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*, 507–537. https://doi.org/10.1111/j.1532-7795.2010.00645.x

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*, 773–794. https://doi.org/10.1037/0022-3514.60.5.773

Goetz, T. (2004). *Emotionales Erleben und selbsreguliertes Lernen bei Schülern im Fach Mathematik* [Students' emotions and self-regulated learning in mathematics]. München: Utz.

Göllner, R., Roberts, B. W., Damian, R. I., Lüdtke, O., Jonkmann, K., & Trautwein, U. (2017). Whose "storm and stress" is it? Parent and child reports of personality development in the transition to early adolescence. *Journal of Personality*, *85*, 376–387. https://doi.org/10.1111/jopy.12246

Göllner, R., Wagner, W., Eccles, J. S., & Trautwein, U. (2018). Students' idiosyncratic perceptions of teaching quality in mathematics: A result of rater tendency alone or an expression of dyadic effects between students and teachers? *Journal of Educational Psychology*, *110*(5), 709–725, https://doi.org/10.1037/edu0000236

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*, 80–100. https://doi.org/10.1207/S15328007SEM1001_4

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102–116. https://doi.org/10.1037/a0038889

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*, 625–638. https://doi.org/10.1111/1467-8624.00301

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. https://doi.org/10.1037//1082-989X.3.4.424

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

Jackson, J. J., Bogg, T., Walton, K. E., Wood, D., Harms, P. D., Lodi-Smith, J., & Roberts, B. W. (2009). Not all conscientiousness scales change alike: A multimethod, multisample study of age differences in the facets of conscientiousness. *Journal of Personality and Social Psychology*, *96*, 446–459. https://doi.org/10.1037/a0014156

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

Johnson, J. A. (1999). Persons in situations: Distinguishing new wine from old wine in new bottles. *European Journal of Personality*, *13*, 443–453. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5xxaaa443:AID-PER358xxbbb3.0.CO;2-9

Jonkmann, K., Rose, N., & Trautwein, U. (Eds.) (2013). *Tradition und innovation: Entwicklungsverläufe an Haupt- und Realschulen in Baden-Württemberg und Mittelschulen in Sachsen – Abschlussbericht für die Länder Baden-Württemberg und Sachsen*. Tübingen: Projektbericht an die Kultusministerien der Länder.

Kandler, C., Zimmermann, J., & McAdams, D. P. (2014). Core and surface characteristics for the description and theory of personality differences and development. *European Journal of Personality*, *28*, 231–243. https://doi.org/10.1002/per.1952

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success*. Cambridge, MA: National Bureau of Economic Research.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*, 27–50. https://doi.org/10.1080/09500693.2010.518645

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen [Test quality and psychometric equivalence of the German version of the Big Five Inventory (BFI) in young, middle-aged and elderly adults]. *Diagnostica*, *47*, 111–121. https://doi.org/10.1026//0012-1924.47.3.111

Leikas, S., Lönnqvist, J.-E., & Verkasalo, M. (2012). Persons, situations, and behaviors: Consistency and variability of different behaviors in four interpersonal situations. *Journal of Personality and Social Psychology*, *103*, 1007–1022. https://doi.org/10.1037/a0030385

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 151–173. https://doi.org/10.1207/S15328007SEM0902_1

Marsh, H. W. (1992). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual*

and a research monograph. Macarthur: University of Western Sydney.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*, 204–217. https://doi.org/10.1037/0003-066X.61.3.204

McCormick, M. P., & O'Connor, E. E. (2015). Teacher–child relationship quality and academic achievement in elementary school: Does gender matter? *Journal of Educational Psychology*, *107*, 502–516. https://doi.org/10.1037/a0037457

McCrae, R. R., & Costa, P. T. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 159–181). New York, NY: Guilford Press.

McCrae, R. R., & Costa, P. T. (2013). Introduction to the empirical and theoretical status of the five-factor model of personality traits. In T. A. Widiger, & P. T. Costa (Eds.), *Personality disorders and the five-factor model of personality* (pp. 15–27). Washington, DC: American Psychological Association. https://doi.org/10.1037/13939-002

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543. https://doi.org/10.1007/BF02294825

Mõttus, R., Allik, J., Hřebíčková, M., Kööts-Ausmees, L., & Realo, A. (2016). Age differences in the variance of personality characteristics. *European Journal of Personality*, *30*, 4–11. https://doi.org/10.1002/per.2036

Mõttus, R., Soto, C. J., & Slobodskaya, H. R. (2017). Are all kids alike? The magnitude of individual differences in personality characteristics tends to increase from early childhood to early adolescence. *European Journal of Personality*, *18*, 313–328. https://doi.org/10.1002/per.2107

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, *54*, 557–585. https://doi.org/10.1007/BF02296397.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th. ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267–316. https://doi.org/10.2307/271070

Organisation for Economic Co-operation and Development. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: Author.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*, 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., Jullien, S., Lichtenfeld, S., Frenzel, A. C., Goetz, T., Vom Hofe, R., & Blum, W. (2005). *Skalenhandbuch PALMA: 4. Messzeitpunkt (8. Klassenstufe) [Scale documentation of the PALMA study: 4th timepoint (grade 8)]*. Munich: Universität München, Institut für Pädagogische Psychologie.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365–386). Boston, MA: Springer. https://doi.org/10.1007/978-1-4614-2018-7_17

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*, 322–338. https://doi.org/10.1037/a0014996

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., … Funder, D. C. (2014). The situational eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, *107*, 677–718. https://doi.org/10.1037/a0037250

Rauthmann, J. F., Sherman, R. A., & Funder, D. C. (2015). Principles of situation research: Towards a better understanding of psychological situations. *European Journal of Personality*, *29*, 363–381. https://doi.org/10.1002/per.1994

Reis, H. T. (2008). Reinvigorating the concept of situation in social psychology. *Personality and Social Psychology Review*, *12*, 311–329. https://doi.org/10.1177/1088868308321721

Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement*. New York, NY: Routledge.

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social cognitive constructs are just as stable as the Big Five between grades 5 and 8. *AERA Open*, *3*, 1–9. https://doi.org/10.1177/2332858417717691

Roberts, B. W. (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality*, *43*, 137–145. https://doi.org/10.1016/j.jrp.2008.12.015

Roberts, B. W. (2018). A revised sociogenomic model of personality traits. *Journal of Personality*, *86*, 23–35. https://doi.org/10.1111/jopy.12323

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*, 31–35. https://doi.org/10.1111/j.1467-8721.2008.00543.x

Roberts, B. W., & Nickel, L. (2017). A critical evaluation of the neo-socioanalytic model of personality. In J. Specht (Ed.), *Personality development across the life span* (pp. 157–177). London: Elsevier.

Roberts, B. W., & Pomerantz, E. M. (2004). On traits, situations, and their integration: A developmental perspective. *Personality and Social Psychology Review*, *8*, 402–416. https://doi.org/10.1207/s15327957pspr0804_5

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*, 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, *75*, 243–248. https://doi.org/10.1007/s11336-009-9135-y

Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology*, *99*, 330–343. https://doi.org/10.1037/a0019796

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718–737. https://doi.org/10.1037/0022-3514.94.4.718

Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence: Structure, development, and outcomes. *Current*

Directions in Psychological Science, *24*, 358–362. https://doi.org/10.1177/0963721415589345

Stamps, J. A., & Biro, P. A. (2016). Personality and individual differences in plasticity. *Current Opinion in Behavioral Sciences*, *12*, 18–23. https://doi.org/10.1016/j.cobeha.2016.08.008

Suls, J., & Mullen, B. (1982). From the cradle to the grave: Comparison and self-evaluation across the life-span. In J. Suls (Ed.), *Psychological perspectives on the self* (pp. 97–125). Hillsdale, NJ: Erlbaum.

Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A., & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of Personality and Social Psychology*, *109*, 142–162. https://doi.org/10.1037/pspp0000034

Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology*, *90*, 202–209. https://doi.org/10.1037/0022-0663.90.2.202

Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, *73*, 287–301. https://doi.org/10.1111/1467-8624.00406

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, *4*, 10–18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, *12*, 265–310. https://doi.org/10.1016/0273-2297(92)90011-P

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81. https://doi.org/10.1006/ceps.1999.1015

How to cite this article: Rieger S, Göllner R, Spengler M, et al. The effects of getting a new teacher on the consistency of personality. *Journal of Personality*. 2019;87:485–500. https://doi.org/10.1111/jopy.12410