AERA Open
July-September 2017, Vol. 3, No. 3, pp. 1–9
DOI: 10.1177/2332858417717691
© The Author(s) 2017. http://ero.sagepub.com

Social Cognitive Constructs Are Just as Stable as the Big Five Between Grades 5 and 8

Sven Rieger
Richard Göllner
Marion Spengler
Ulrich Trautwein
Benjamin Nagengast
University of Tübingen
Brent W. Roberts
University of Illinois at Urbana-Champaign and University of Tübingen

Personality traits and social cognitive variables are central constructs in psychological research. It is often assumed that personality traits are less changeable than social cognitive variables, and thus interventions usually tend to focus on the latter. However, these assumptions about the mutability of personality and social cognitive variables have never been tested side-by-side in a longitudinal study. Using a large longitudinal study with four time points and N *= 3,876 students in 136 classes (99 schools), we compared the mutability of the Big Five personality traits and social cognitive constructs more often described as socioemotional skills or motivational variables (individual interest, self-concept, and academic effort in the school subjects math, English, and German). The results indicated that there are no marked differences between personality traits and social cognitive constructs across multiple indicators of stability or changeability.*

Keywords: *Big Five, social cognitive constructs, stability, latent-state-trait models*

Over the past decade there has been a groundswell of attention brought to bear on "noncognitive" factors that contribute to school outcomes. Motivated in part by research emerging out of economics that showed that factors other than cognitive ability were important for human capital (Heckman, Stixrud, & Urzua, 2006), researchers and policymakers from multiple fields, but especially education, have been investigating the importance of noncognitive factors. Multiple reviews have come to the same conclusion: Noncognitive factors are critical for success in both educational and occupational settings, and interventions should focus on these factors (Kautz, Heckman, Diris, ter Weel, & Borghans, 2014).

What are noncognitive factors? As the name would indicate, noncognitive factors are any constructs that are not considered traditional indicators of cognitive ability or intellectual functioning. From the viewpoint of psychology, this broad category inevitably includes a mix of many different constructs that go by many different names and that emerge out of disparate theoretical orientations. Common terms used to describe noncognitive factors are *socioemotional skills*, *character*, *personality*, and *21st-century skills*. Prototypical constructs include factors such as self-concept of ability, self-efficacy, academic persistence, conscientiousness, stress tolerance, grit, and creativity (National Research Council, 2012).

The rather inclusive grouping of concepts belies the often stark theoretical and conceptual distinctions dividing these constructs when used in research. Some of these variables, represented best by conscientiousness, are considered *traits* and are often described as enduring, broad, and heritable (McCrae & Costa, 2008). On the other side are social cognitive variables (e.g., self-efficacy), which are presumed to be narrow, relevant to very specific contexts, and derived almost exclusively from experience rather than genetics (Bandura, 2012). The distinction between these two groups of variables is supported by many different theoretical models that conceptualize traits as core characteristics and social cognitive variables as surface characteristics (Asendorpf & van Aken, 2003). The fact that these conceptually distinct constructs are brought together into one group reflects, in part, the pragmatism of many educational researchers and economists who have discovered that all of these constructs appear to be useful for predicting important educational and human-capital outcomes (Almlund, Duckworth, Heckman, & Kautz, 2011).

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Given their importance to policy-relevant outcomes (e.g., educational attainment and occupational success), the discussion has quickly turned to how these qualities can be fostered in students through interventions, especially in childhood and adolescence (Heckman & Kautz, 2012).[1] For a noncognitive quality to be a viable target of an intervention, it is typically assumed that it should be malleable. Currently, the prevailing belief is that the malleability of a noncognitive characteristic is perfectly aligned with the conceptual distinctions that have been drawn between the trait and social cognitive perspectives. Social cognitive variables (e.g., math self-efficacy or interest in social science) are often assumed to be more malleable than constructs associated with personality (e.g., grit and conscientiousness; Bailey, Duncan, Odgers, & Yu, 2015; Harter, 1998; Shavelson, Hubner, & Stanton, 1976). Based on the prevailing theoretical systems in psychology, this assumption appears well justified when considering that conceptually, social cognitive constructs emerged from a framework that by definition presents most of its affiliated constructs as malleable (Bandura, 2012). In particular, social cognitive variables are traditionally seen as more entwined with individuals' social contexts. For example, self-concept beliefs emerged as a result of social comparison processes and the evaluation of one's own ability (Suls & Mullen, 1982). Also, domain-specific interests were found to result from a student's interaction with a specific object or activity (Krapp, 2002) and are presumed not to be based on biological factors (but see Kovas et al., 2015). By virtue of their theoretical origins, social cognitive constructs are assumed to be amenable to change, and there is little doubt that they can be fostered through interventions (see, e.g., Lazowski & Hulleman, 2016; O'Mara, Marsh, Craven, & Debus, 2006).

In contrast, there is some debate about the ratio of stability and change in personality traits (Anusic & Schimmack, 2016). Personality traits are often assumed to be highly heritable and highly stable and therefore not amenable to change. For instance, the five-factor theory (McCrae & Costa, 2013) asserts that personality traits are influenced primarily by biological factors such as genetic predispositions, and trait change is attributed solely to intrinsic maturation and not to life experience or environmental effects. In contrast, research emerging from the Neo-socioanalytic framework of personality holds that personality traits are not perfectly stable and can be affected by experience (Roberts & Nickel, 2017). According to this perspective, personality traits in particular are marked by relative stability and change (both mean-level and individual differences) that are presumed to be caused by experiences and environmental factors (Caspi, Roberts, & Shiner, 2005; Roberts, Walton, & Viechtbauer, 2006).

Whereas the assumptions of the relative malleability of social cognitive and personality constructs appear reasonable, data that support these positions are surprisingly scarce. Although it is common to intervene or try to change social cognitive constructs such as socioemotional skills (e.g., Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011), it is less common to focus on their test–retest stability or heritability. Conversely, although the idea that personality traits do develop with time and age is beginning to gain a foothold in the personality literature (e.g., Roberts & Mroczek, 2008), it is highly unusual to find research focused on intervening to change personality traits, and conversely much more common to find reports of their test–retest stability and heritability (e.g., Ferguson, 2010). When examined separately in various longitudinal studies, it is common to find that mean-level changes in both personality traits and social cognitive constructs in adolescence are surprisingly heterogeneous (Musu-Gillette, Wigfield, Harring, & Eccles, 2015; Soto, John, Gosling, & Potter, 2011; Van den Akker, Deković, Asscher, & Prinzie, 2014). Fittingly, a recent review of the conceptual and empirical basis of the distinction between these two classes of constructs found that the division between traits and social cognitive concepts was conceptually larger than the empirical data would justify (Kandler, Zimmermann, & McAdams, 2014).

Nonetheless, because the intellectual heritage behind these two classes of variables is so stark, it is uncommon to find both types of noncognitive constructs included in the same study (Roberts, 2009). This has created an asymmetry in the understanding of how changeable constructs like social cognitive variables and personality traits may be. In the absence of systematic interventions on both social cognitive variables and personality traits, researchers are left with observational data as the basis to inferences about their relative consistency and mutability.

The Present Study

Although not directly addressing whether a concept can be changed through intervention, passive observational studies can provide valuable information on the continuity and change of concepts over time and thus their potential for changeability. Presumably, if one class of variable would be more consistent, and in turn show less change over time in the same longitudinal study, this would lend credence to the argument that constructs like conscientiousness are, or are not, good targets for intervention. The problem is that, to our knowledge, no study has explicitly examined the stability and mutability of both sets of constructs within the same longitudinal sample of students. A focus on students is key, as they make up the populations that are the focus of most interventions, under the presumption that the investments made in these populations will reap larger gains throughout life (Cunha & Heckman, 2010). In an effort to address this oversight, we contrast social cognitive variables and personality traits in terms of their temporal continuity and change over time within a large scale, longitudinal study of students with four time points (N = 3,876 in 136 classes; age range

11–14). To this end, Big Five personality traits of extraversion, agreeableness, conscientiousness, neuroticism, and openness were examined as well as social cognitive constructs such as subject-specific individual interest, self-concept, and academic effort. Because continuity and mutability are not unitary constructs, we examined multiple indices of continuity and change over time to determine whether these two classes of variables differed in the baseline consistency or mutability. Based on test–retest correlations and parameters from generalized second-order growth models (GSGM; Bishop, Geiser, & Cole, 2015), three research questions were examined. First, how stable are the constructs over time? Second, to what degree is the stability of each class of constructs attributable to unchanging components, and how much of the instability is attributable to state components? Third, are interventions predicated on being able to move individuals or groups of individuals in one direction or another? One way to characterize the malleability of a noncognitive characteristic is therefore to examine the extent to which people show increases or decreases in the constructs, which is typically examined as individual differences in change (Roberts & Mroczek, 2008). We examine the average level of individual differences in change across these two classes of constructs to get a better idea of whether individuals change more or less on each class of variables in a naturalistic longitudinal study.

Method

Sample

We used data from a large longitudinal German study, Tradition and Innovation (TRAIN), which is hosted by the Hector Research Institute of Education Sciences and Psychology. TRAIN is a large-scale school achievement study that encompasses four time points (T1, T2, T3, and T4) from Grades 5 to 8. The study comprises 136 classes in 99 schools from two federal states (Baden-Württemberg and Saxony). Data were available for n = 2,894 (46% female) individuals at T1 (Grade 5), n = 2,936 (45% female) individuals at T2 (Grade 6), n = 2,993 (46% female) individuals at T3 (Grade 7), and n = 3,060 (46% female) individuals at T4 (Grade 8). The sample size of the pooled data set was N = 3,876. This data set contained all individuals who gave information at a minimum of one time point.[2]

Instruments

The social cognitive constructs (i.e., self-concept, interest, and academic effort) were assessed with four items per school subject (math, German, English). 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 self-concept items targeted the students' own evaluation of their ability in the respective school subjects. The domain-specific interest items focused on the intrinsic value of the respective school subject. The items from the academic effort scales were focused on the effort needed to meet subject-specific tasks. 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*).[3] Unfortunately, the negatively worded items showed negative or low item-total correlations for all Big Five traits (all $rs < .22$). Therefore, we used only the positively worded items. We conducted several robustness checks in examining the item properties of all Big Five scales as a whole by means of confirmatory factor analyses. Following this, we reran all of our analyses with the complete set of items (for the results see online Supporting Information [SI] Appendix C). The main results remained virtually unchanged. Sample items, the number of items, and the Cronbach's alpha values (ranging from .55 to .90) at all four time points are shown in Table S3 in the SI Appendix A2.

Statistical Analyses

All models were estimated in the framework of longitudinal confirmatory factor analyses in Mplus 7.31 (Muthén & Muthén, 1998–2012). Two-sided statistical tests were performed at the 5% level of significance.

The analytical procedure encompassed roughly three steps. First, we estimated manifest (rank-order) correlations between all constructs; from these models, means, standard deviations, as well as the rank-order correlations of the adjacent time points were derived. Second, to properly interpret latent variable change in longitudinal models, at least strong measurement invariance has to be established (Meredith, 1993; Meredith & Teresi, 2006); thus, we specified latent-state models in which we imposed strong measurement invariance (same loadings and intercepts for each indicator over time), separately for each construct. Third, we estimated GSGMs to model the function of mean-level change and, in addition, to separate the rather stable variance from the time-point-specific variance (Bishop et al., 2015). The GSGM is a combination of a latent-state-trait (LST) model and a latent growth curve (LGC) model (Geiser et al., 2015) and is diagrammed in Figure S1 in the SI Appendix A3.

In LST models three coefficients are of particular interest: the consistency (CO), occasion specificity (OS), and reliability (REL) coefficients (Steyer, Schmitt, & Eid, 1999). The CO coefficients indicate the degree of stability across measurement points (trait part). The OS coefficients denote the degree to which differences in observed variables are influenced by the situation or by Person × Situation interactions (state part). The REL coefficients reveal the extent to which the measures are reliable and are not due to measurement error.

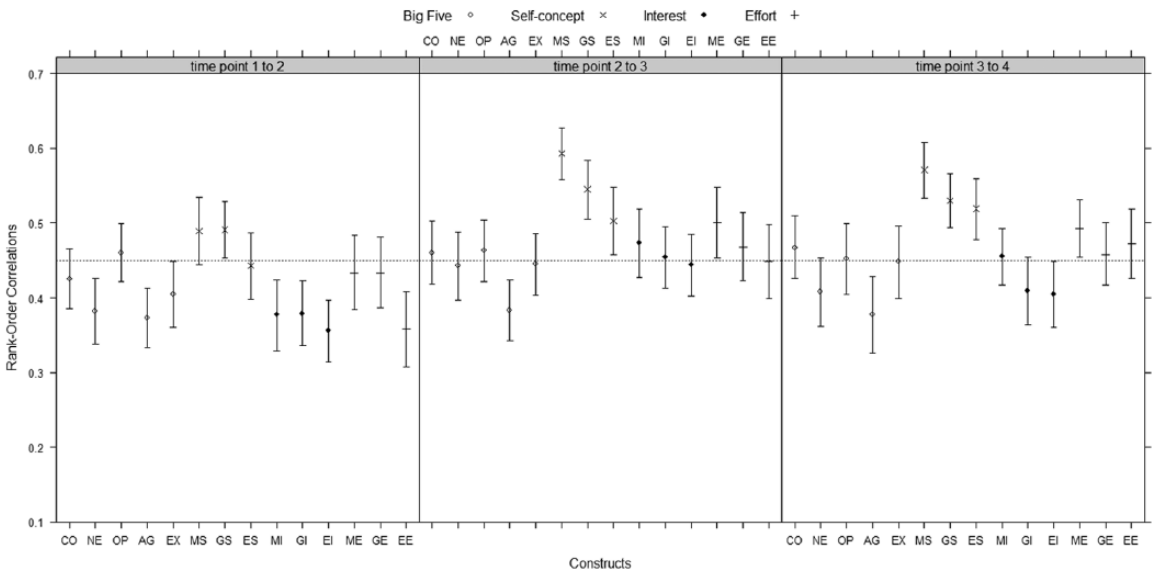

FIGURE 1. *Manifest rank-order correlations with 95% confidence intervals. CO = conscientiousness; NE = neuroticism; AG = agreeableness; EX = extraversion; MS = math self-concept; GS = German self-concept; ES = English self-concept; MI = math interest; GI = German interest; EI = English interest; ME = math effort; GE = German effort; EE = English effort. The dashed line indicates the mean rank-order correlation for all constructs ($\bar{r}$ = .45).*

Missing data. Due to attrition or nonresponse at single time points, missing values occur in nearly all longitudinal studies. To deal with these, we used the full-information maximum likelihood procedure (see, e.g., Enders, 2001). Due to its ability to offer less-biased parameter estimates, this procedure is believed to be superior to conventional methods, such as listwise or pairwise deletion (Graham, 2009). Furthermore, to make the missing at random assumption more plausible, we included several auxiliary variables (Collins, Schafer, & Kam, 2001) in all analyses (e.g., standardized achievement tests, grades, socioeconomic status, gender, etc.).

Nested data structure. In the present study, students were nested within classes, resulting in a multilevel structure in the data set. Therefore, students in a class are not independent from each other (students within classes tend to be more similar than students from different classes; Raudenbush & Bryk, 2002). Failing to consider this structure could lead to an underestimation of the standard errors (see e.g., Muthen & Satorra, 1995). However, we were not interested in contextual effects and thus relied on single-level analyses with cluster-robust standard errors (McNeish, Stapleton, & Silverman, 2016) as implemented in Mplus. In addition, it should be noted that the intraclass correlation coefficients were rather small for the Big Five and social cognitive constructs, ranging from .02 to .09.

Results

To contrast personality traits and social cognitive constructs in terms of their temporal stability over time, we began by computing simple descriptive statistics and rank-order correlations. Table S4 in the SI Appendix B1 presents the means and standard deviations of all constructs. The means of two Big Five measures (conscientiousness, openness) and all social cognitive constructs (except self-concept in the subject German) consistently decreased over time. Neuroticism, extraversion, and agreeableness as well as the self-concept in the subject German did not increase or decrease over time. The standard deviations were slightly smaller for the Big Five measures than for the social cognitive constructs. The rank-order correlations between adjacent time points for all constructs are depicted in Figure 1. In addition, the average rank-order correlation ($\bar{r}$ = .45) of all constructs is depicted as a reference point. For the Big Five, the rank-order correlations ranged from .37 to .47 and were usually not statistically significantly different from the average rank-order correlation. The highest rank-order correlations were found for domain-specific self-concept measures ranging from .44 to .59. With the exception of the rank-order correlations from T1 to T2 for math and English, the rank-order correlations were statistically significantly higher than the average rank-order correlation. For subject-specific interest, the rank-order correlations ranged from .36 to .47. For academic effort, they ranged from .36 to .50, and they were usually not statistically different from the average rank-order correlation (see also Table S5 in the SI Appendix B2). In addition, we compared the average rank-order correlation of the Big Five traits ($\bar{r}$ = .43) with the average rank-order correlation of the social cognitive constructs ($\bar{r}$ = .46). The average rank-order correlation of the social cognitive constructs was statistically significantly higher than

TABLE 1
Results of the Generalized Second-Order Growth Models

Construct	Intercept	Intercept variance	Slope	Slope variance	COR (I, S)
Big Five					
Conscientiousness	**2.93**	**0.14**	**−0.05**	**0.016**	**−.52**
	[2.89, 2.96]	**[0.12, 0.16]**	**[−0.06, −0.04]**	**[0.012, 0.020]**	**[−.60, −.45]**
Neuroticism	**2.55**	**0.20**	**−0.02**	**0.017**	**−.51**
	[2.51, 2.59]	**[0.14, 0.25]**	**[−0.03, −0.01]**	**[0.008, 0.027]**	**[−.63, −.40]**
Openness	**2.89**	**0.18**	**−0.04**	**0.018**	**−.53**
	[2.86, 2.92]	**[0.15, 0.21]**	**[−0.05, −0.03]**	**[0.012, 0.023]**	**[−.61, −.45]**
Agreeableness	**2.80**	**0.12**	**−0.01**	**0.013**	**−.56**
	[2.76, 2.83]	**[0.08, 0.15]**	**[−0.02, −0.00]**	**[0.008, 0.018]**	**[−.67, −.45]**
Extraversion	**2.92**	**0.17**	0.00	**0.020**	**−.54**
	[2.88, 2.95]	**[0.11, 0.23]**	[−0.01, 0.02]	**[0.010, 0.030]**	**[−.63, −.44]**
Social cognitive constructs					
Self-concept in math	**3.03**	**0.30**	**−0.09**	**0.027**	**−.25**
	[2.98, 3.07]	**[0.25, 0.35]**	**[−0.11, −0.07]**	**[0.018, 0.037]**	**[−.38, −.12]**
Self-concept in German	**3.01**	**0.14**	−0.01	**0.007**	**−.36**
	[2.98, 3.04]	**[0.10, 0.17]**	[−0.02, 0.00]	**[0.003, 0.012]**	**[−.51, −.21]**
Self-concept in English	**3.05**	**0.19**	**−0.07**	**0.023**	**−.31**
	[3.01, 3.08]	**[0.16, 0.23]**	**[−0.08, −0.06]**	**[0.016, 0.030]**	**[−.44, −.18]**
Interest in math	**3.11**	**0.11**	**−0.12**	**0.008**	−.06
	[3.07, 3.14]	**[0.08, 0.14]**	**[−0.14, −0.11]**	**[0.003, 0.014]**	[−.36, .24]
Interest in German	**3.03**	**0.18**	**−0.12**	**0.011**	**−.33**
	[2.98, 3.07]	**[0.14, 0.22]**	**[−0.13, −0.10]**	**[0.004, 0.019]**	**[−.53, −.13]**
Interest in English	**3.07**	**0.15**	**−0.12**	**0.010**	−.18
	[3.03, 3.12]	**[0.11, 0.18]**	**[−0.14, −0.11]**	**[0.002, 0.017]**	[−.47, .11]
Effort in math	**3.27**	**0.17**	**−0.11**	**0.018**	−.11
	[3.24, 3.30]	**[0.13, 0.20]**	**[−0.12, −0.09]**	**[0.011, 0.025]**	[−.29, .08]
Effort in German	**3.28**	**0.16**	**−0.08**	**0.016**	**−.23**
	[3.24, 3.31]	**[0.13, 0.19]**	**[−0.09, −0.07]**	**[0.010, 0.023]**	**[−.38, −.09]**
Effort in English	**3.33**	**0.13**	**−0.09**	**0.017**	−.12
	[3.29, 3.36]	**[0.10, 0.16]**	**[−0.11, −0.08]**	**[0.011, 0.024]**	[−.30, .06]
Average coefficients for the Big Five	2.82	0.16	−0.02	0.017	−.53
Average coefficients for the social cognitive constructs	3.13	0.17	−0.09	0.015	−.21
Overall average coefficients	3.02	0.17	−0.07	0.016	−.33

Note. N = 3,876. Coefficients in bold are statistically significantly different from 0 ($p < .05$, two tailed). The values in brackets are 95% confidence intervals. COR (I, S) = correlation between intercept and slope.

the one for the Big Five traits ($Z = 3.50$, $p = .000$). However, the difference was rather small (Δ .037).

With respect to the second and third research questions (degree of stable and time-point-specific stability or change and average level of individual differences in change), we estimated GSGM separately for each construct. For these analyses, we assumed linear change. The assumption appeared to be reasonable, because the means decreased consistently over time (see again Table S4 in the SI Appendix B1), and all models demonstrated good fit (see Table S6 in the SI Appendix B3). Table 1 presents the intercepts, slopes, and their variances as well as the correlation between the intercepts and slopes. In addition, we again provide the average coefficients of all parameters as reference points. The statistically significant negative slope factors (except for self-concept in the subject German and extraversion) coincided with the decreasing means over time found in the manifest indicators. Furthermore, the GSGM models allowed us to distinguish between the stable variance components (CO), state variability (time-point-specific variance components; OS), and measurement error (ERR).[4] The different variance proportions of the constructs are depicted in Figure 2.[5] We again provide the mean coefficients as reference points (see Table S8 in SI Appendix B5). For the Big Five traits, the stable variance components captured 30% to 42% of the variance, whereas 21% to 28% of the variance was time-point-specific, and 30% to 47% was due to

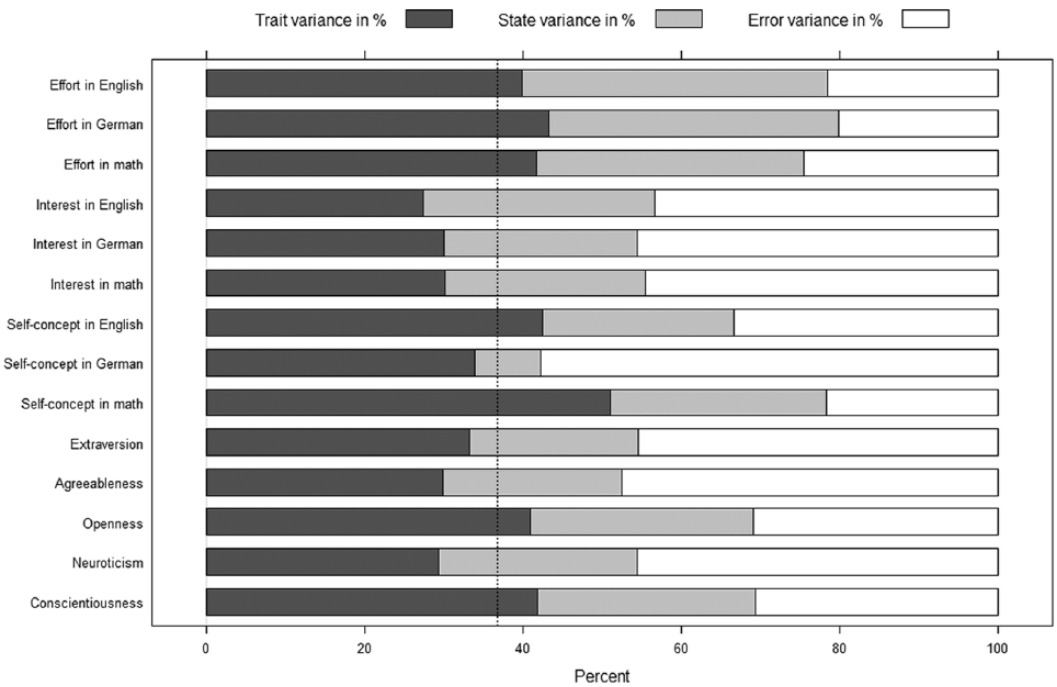

FIGURE 2. *Proportions of variance in percentages for all constructs. The dashed line indicates the mean proportion of trait variance across all constructs (36.8%).*

measurement error. For the social cognitive constructs, 27% to 51% of the variance was explained by the CO, whereas 8% to 39% was due to state variability, and 22% to 53% was due to measurement error. Attention should be drawn to the measure of German self-concept. In comparison with the other constructs, it showed a small amount of state variance and a large amount of measurement error. Examining the trait variance components for the Big Five, we see that conscientiousness (Δ 5.0%, $SE = 0.016$, $p = .002$) and openness (Δ 4.1%, $SE = 0.014$, $p = .005$) were statistically significantly above the average reference point (36.8%). For the social cognitive variables, self-concept in math (Δ 14.2%, $SE = 0.015$, $p < .000$) and English (Δ 5.7%, $SE = 0.015$, $p < .000$) as well as academic effort in math (Δ 4.9%, $SE = 0.015$, $p = .001$) and German (Δ 6.5%, $SE = 0.020$, $p = .001$) were statistically significantly above the average reference point. Statistically significantly below the average reference point were neuroticism (Δ –7.5%, $SE = 0.017$, $p < .000$), agreeableness (Δ –7.0%, $SE = 0.022$, $p = .002$), and extraversion (Δ –3.6%, $SE = 0.016$, $p = .022$) for the Big Five and the interest measures in math (Δ –6.7%, $SE = 0.014$, $p < .000$), German (Δ –6.8%, $SE = 0.015$, $p < .000$), and English (Δ –9.4%, $SE = 0.013$, $p < .000$) for the social cognitive variables. Self-concept in German and academic effort in English were not statistically significantly different from the reference point. In sum, the same number of scales was above and below the average amounts of stable, state, and error variance for both the Big Five and the social cognitive variables, showing no evidence for differential stability or changeability across construct domains.

As a last step, we compared the variances in the slope parameters for each construct, as this represents the average amount of individual differences in change across the social cognitive and personality trait indicators. The variances in the slopes provide an estimate of the average range of change that students exhibited in each class of variables over time. We compared all slope variances of the constructs against the average (.016). Except for the self-concept measures in math (Δ 0.011, $SE = 0.005$, $p = .016$), German (Δ –0.007, $SE = 0.004$, $p = .041$), and English (Δ 0.007, $SE = 0.004$, $p = .041$), as well as interest in math (Δ –0.008, $SE = 0.003$, $p = .005$), we found that no slope variance was statistically significantly different from the average slope variance across the Big Five or the social cognitive domains.

Discussion

In the present study, we examined the stability and mutability of the Big Five personality traits and social cognitive constructs such as subject-specific individual interest, self-concept, and academic effort in the same longitudinal sample of students. Our results showed no meaningful differences in (a) the stability of the constructs at the manifest level, (b) the percentage of stable variance for each construct, or (c) the amount of change that students showed on each construct over time in the form of the average variance in the slopes. Based on this study, we found no meaningful differences between the stability and mutability of social cognitive or personality trait variables in a longitudinal sample of

adolescent students. However, it should be noted that there were heterogeneous results within each construct class. For instance, within the Big Five traits, conscientiousness and openness were more stable than extraversion, neuroticism and agreeableness. On the other hand, within the social cognitive constructs, the self-concept measures were more stable than the interest measures.

The findings of this study are important for educational policymakers who are currently focused on how to best enhance noncognitive skills (Heckman & Kautz, 2012) as an adjunct to focusing exclusively on cognitive skills. In this context, it is important to understand the nature of the various constructs that fall under the umbrella of noncognitive constructs, as their scope and variety are more diverse than those typically found in cognitive skills.

It is also the case that there are many assumptions behind the different classes of constructs within the noncognitive set, with the most prevalent being that personality traits are not changeable and social cognitive constructs are changeable and therefore the latter should be the focus of interventions (Bailey et al., 2015). Whereas our study could not directly address whether either class of variable could be more easily changed through a direct intervention, our study tests basic assumptions behind the stability and mutability of these classes of constructs. Succinctly, many parties assume that personality traits, even in childhood, represent stable, unchangeable constructs (Bailey et al., 2015). By contrast, social cognitive and motivational constructs are typically considered to be less stable and more changeable. Up to this point, the relative stability and mutability of these constructs has never been tested simultaneously in a longitudinal study of students, the population most often considered for intervention. Our study provides valuable data on this comparison and shows that both classes of variables are equivalently stable and equivalently changeable in an observational longitudinal study. Whereas this does not prove that personality traits are changeable through intervention or that social cognitive variables are difficult to change, it does warrant a more open approach to considering which variables to include in future intervention studies.

Limitations and Future Directions

Although the current study used a large sample over a 3-year period of time and examined a variety of social cognitive and personality constructs, some noteworthy limitations should be kept in mind when interpreting the results. First, it should be noted that students from the highest school track in Germany (i.e., the Gymnasium) did not participate in the study. Rather, the students came from the other nonacademic tracks (lower and intermediate track schools) in two states of Germany. Thus, it would be ideal for future studies to include students from all academic tracks. Second, we used only the positively worded items from the Big Five scales because the negatively worded items demonstrated low, zero, or even negative item-total correlations. These results probably had to do with response biases in terms of acquiescent tendencies, midpoint, or extreme responding and made it very challenging to analyze these measures. This issue is being discussed in the literature on using self-reports in young children and clearly needs further consideration (see, e.g., Soto, John, Gosling, & Potter, 2008). Third, analyzing whether there are comparable stabilities and trait-state ratios for other instruments would be indispensable for obtaining a more thorough understanding of the relative stability of these two classes of constructs. Finally, there is a need to examine the trait-state proportions on the facet level of the personality traits as well as other classes of person characteristics (e.g., motives, abilities, etc.).

Acknowledgments

This research was supported by the Excellence Initiative of the German federal and state governments through the LEAD Graduate School and Research Network; by a grant to Marion Spengler funded by the European Social Fund and the Ministry of Science, Research and the Arts of Baden-Württemberg; and by a grant to Ulrich Trautwein from the Ministry of Science, Research and the Arts of Baden-Württemberg (Az: 33-7532.20/735). We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen.

Notes

1. The periods of late childhood and early adolescence are defined by fundamental changes in youths' lives (e.g., rapid biological changes, shifting demands in school life, initiation of new relationships with peers, etc.; Soto & Tackett, 2015) that are potential sources of instability and also mutability. These age periods are ideal for studying these issues as this is when noncognitive factors are the most relevant to educational outcomes and are most often the targets of interventions.

2. Supporting information on the sample composition is reported in the online Supporting Information (SI) Appendix A1.

3. We applied a linear transformation (scale = 3/4 + 0.25) to convert the 5-point Likert scale to a scale with a minimum of 1 and a maximum of 4 so that we could compare it with the motivational constructs.

4. Detailed information on the consistency (CO), occasion specificity (OS), and error (ERR) coefficients is described in the SI Appendix A3.

5. This is a simplified presentation of the results. In latent-state-trait (LST) models, each indicator (parcel) has its own consistency, specificity, and reliability coefficient. To simplify matters, the consistency and specificity coefficients were aggregated across the item parcels and time points. The specific results for the item parcels as well as the averaged LST coefficients are reported in Tables S7 and S8 in the SI Appendices B4 and B5.

References

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S.

Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp. 1–181). Amsterdam, Netherlands: North Holland. doi.org/10.1016/B978-0-444-53444-6.00001-8

Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology*, *110*, 766–781. doi.org/10.1037/pspp0000066

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. doi.org/10.1111/1467-6494.7104005

Bailey, D., Duncan, G. J., Odgers, C., & Yu, W. (2016). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*, 7–39.

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

Bishop, J., Geiser, C., & Cole, D. A. (2015). Modeling latent growth with multiple indicators: A comparison of three approaches. *Psychological Methods*, *20*, 43–62. doi.org/10.1037/met0000018

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

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. doi.org/10.1037/1082-989X.6.4.330

Cunha, F., & Heckman, J. (2010). *Investing in our young people*. Cambridge, MA: National Bureau of Economic Research.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432. doi.org/10.1111/j.1467-8624.2010.01564.x

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, *6*, 352–370.

Ferguson, C. J. (2010). A meta-analysis of normal and disordered personality across the life span. *Journal of Personality and Social Psychology*, *98*, 659–667. doi.org/10.1037/a0018770

Geiser, C., Keller, B. T., Lockhart, G., Eid, M., Cole, D. A., & Koch, T. (2015). Distinguishing state variability from trait change in longitudinal data: The role of measurement (non) invariance in latent state–trait analyses. *Behavior Research Methods*, *47*, 172–203. doi.org/10.3758/s13428-014-0457-z

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576. doi.org/10.1146/annurev.psych.58.110405.085530

Harter, S. (1998). Developmental perspectives on the self-system. In N. Eisenberg (Ed.), *Handbook of child psychology* (pp. 553–618). New York, NY: Wiley.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*, 451–464. doi.org/10.1016/j.labeco.2012.05.014

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*, 411–482. doi.org/10.1086/504455

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

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. doi.org/10.1002/per.1952

Kautz, T., Heckman, J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (NBER Working Paper No. 20749). Cambridge, MA: National Bureau of Economic Research.

Kovas, Y., Garon-Carrier, G., Boivin, M., Petrill, S. A., Plomin, R., Malykh, S. B., . . . Vitaro, F. (2015). Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *Personality and Individual Differences*, *80*, 51–63. doi.org/10.1016/j.paid.2015.02.006

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*, 383–409. doi.org/10.1016/S0959-4752(01)00011-1

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. doi.org/10.1026//0012-1924.47.3.111

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*, 602–640. doi.org/10.3102/0034654315617832

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.

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. doi.org/10.1037/13939-002

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2016). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*, 114–140. doi.org/10.1037/met0000078

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care*, *44*, 69–77. doi.org/10.1097/01.mlr.0000245438.73837.89

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation*, *21*, 343–370. doi.org/10.1080/13803611.2015.1057161

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

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

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press. doi.org/10.17226/13398

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist*, *41*, 181–206. doi.org/10.1207/s15326985ep4103_4

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

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

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*, 31–35. 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., 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. doi.org/10.1037/0033-2909.132.1.1

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407–441. doi.org/10.3102/00346543046003407

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. doi.org/10.1037/0022-3514.94.4.718

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*, 330–348. doi.org/10.1037/a0021717

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. doi.org/10.1177/0963721415589345

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*, 389–408. doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

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* (1st ed., pp. 97–125). Hillsdale, NJ: Lawrence Erlbaum.

Van den Akker, A. L., Deković, M., Asscher, J., & Prinzie, P. (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting. *Journal of Personality and Social Psychology*, *107*, 736–750. doi.org/10.1037/a0037248

Authors

SVEN RIEGER is a doctoral student at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen and is associated with the LEAD Graduate School and Research Network. He is interested in the development and interplay of personality and motivational variables.

RICHARD GÖLLNER is a senior researcher and junior research group leader at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. His main research interests concern the mental well-being of children and adolescents, personality development, and academic success within the educational context.

MARION SPENGLER is a senior researcher at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. Her research interests are located at the interface of personality psychology and education research.

ULRICH TRAUTWEIN is a professor at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. His research interests include the development of student motivation, personality, academic effort, and achievement.

BENJAMIN NAGENGAST is a professor at the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen. His research interests include quantitative methods (causal inference, latent variable models, multilevel modeling), motivation and academic self-concept, educational effectiveness, and the evaluation of educational interventions.

BRENT W. ROBERTS is a professor of psychology in the Social-Personality-Organizational Division of the Department of Psychology at the University of Illinois and is Distinguished International Professor at the LEAD Graduate School and Research Network. He is interested in the patterns of continuity and change in personality across the decades of adulthood and the mechanisms that affect these patterns.