

Measuring Grit

A German Validation and a Domain-Specific Approach to Grit

Fabian T. C. Schmidt,¹ Johanna Fleckenstein,¹ Jan Retelsdorf,¹ Lauren Eskreis-Winkler,² and Jens Möller³

¹IPN – Leibniz Institute for Science and Mathematics Education, Kiel, Germany

²Department of Psychology, University of Pennsylvania, USA

³IPL – Institute of Psychological Learning Research, University of Kiel, Germany

Abstract: The construct *grit* originates from positive psychology and describes an individual's tendency to persistently pursue long-term goals despite challenges or obstacles. Previous research has shown that domain-general grit is a predictor of educational and vocational success. The present research aimed to establish and validate a German version of the Short Grit Scale by Duckworth and Quinn (2009), named the BISS-8 (*B*eharrlichkeit and *B*eständiges *I*nteresse) Scale, and to test for the domain specificity of grit in an educational context. We conducted three studies to investigate the BISS-8 Scale: in Study 1 (N = 525 university students) confirmatory factor analyses (CFAs) replicated a two-dimensional higher-order structure for the scale. Study 2 (N = 173 university students) investigated the correlations of grit with external criteria such as grade point average (GPA), generalized self-efficacy, general academic self-concept, and personality traits. Finally, in Study 3 (N = 271 high school students), we found differential correlations with school achievement for domain-specific grit. Moreover, the validity of the BISS-8 Scale was also supported for adolescents by replicating the measurement model. All in all, our results indicate the validity of the BISS-8 Scale and show the importance to account for grit in different domains.

Keywords: grit, perseverance, BISS-8 Scale, validation, domain-specific grit

Grit is a construct originating in positive psychology that can be defined as trait-level perseverance of effort and consistency of interest (Duckworth, Peterson, Matthews, & Kelly, 2007). It entails "working strenuously toward challenges, maintaining effort and interest over years despite failure, adversity, and plateaus in progress" (Duckworth et al., 2007, pp. 1087-1088). The construct was introduced by Duckworth et al. (2007) in light of a history of research investigating the relevant indicators for personal success. Based on biographical information, abilities such as perseverance, zeal, and integration toward goals seemed to be important for achievement in addition to talent or intelligence (Cox, 1926; Galton, 1892; Terman & Oden, 1947). Duckworth et al. (2007) integrated these findings by introducing a concept that taps these traits and narrows them down to two dimensions: perseverance of effort and consistency of interest. To achieve a certain level of mastery, deliberate practice and initial failures have to be tolerated and, while this process can take months or even years, the initial interest needs to be kept salient.

To measure grit, Duckworth et al. (2007) developed and validated a self-report questionnaire based on the construct definition of grit as described above. The scale was intended to be used for adolescents and adults who pursue goals in a variety of domains, thus as a domain-general measure (e.g., not just work or school; Duckworth et al., 2007). On these grounds, the authors selected and piloted items for their measure in the US, resulting in the 12-Item Grit Scale. They identified a two-factor structure for the scale. The model fit for a two-factor solution suggested room for improvement, but the scale showed high internal consistency and predictive validity for a variety of success outcomes. In a subsequent validation study, Duckworth and Ouinn (2009) revisited the issue of model fit. The model fit improved substantially when two items were removed from each subscale, leaving eight items in total. The resulting 8-Item scale was named the Short Grit Scale (Grit-S). A second-order factor was included in the model, representing the superordinate latent construct grit. Even though they found differential associations with predicted outcomes for the two subscales, their pattern of findings supports the conceptualization of grit as a compound trait.

To our knowledge, only few studies have been published on grit in a European context so far (Akın & Arslan, 2014; Arslan, Akın, & Çitemel, 2013; Dumfart & Neubauer, 2016) and a validation study for measures of grit in German does not exist. For a fruitful and critical appraisal in the Germanspeaking scientific community, a validation of the German scale seems necessary.

437

This paper aims to provide evidence for the factorial structure as well as the validity of the German short form of the Grit Scale (Grit-S; Duckworth & Quinn, 2009). First, this will be done by replicating the factor structure, testing measurement invariance, and investigating the relationships with external criteria. Second, we will present findings on an issue that has been discussed in previous research (Duckworth & Quinn, 2009), but has not been attended to empirically yet: the domain specificity of grit. While grit has been shown to be relevant for success in several domains, it has not been investigated whether grit itself can be considered a domain-specific construct. Before we present our research in more detail, we will provide a short summary of the state of research on grit.

During the last decade a substantial body of research has found grit to be a relevant predictor of academic and vocational outcomes. Grit seems to explain additional variance in the investigated outcomes over and above intelligence, domain-specific skills, and other personality traits (Duckworth et al., 2007; Duckworth & Seligman, 2006; Duckworth, Tsukayama, & Geier, 2010; Zimmerman & Schunk, 2011). Further research by Duckworth and Quinn (2009) showed that grit was associated with fewer career changes and higher educational attainment. Results of a longitudinal study indicated that adolescents' grit was positively associated with grade point average (GPA) and negatively associated with hours watching television. In addition, grit predicted retention rates for cadets at the United States Military Academy as well as final round attainment for Scripps National Spelling Bee competitors (Duckworth & Quinn, 2009). Moreover, Duckworth, Quinn, and Seligman (2009) found that grittier novice teachers had a more positive effect on the learning outcomes of their students than less gritty teachers. Finally, Robertson-Kraft and Duckworth (2014) showed that grittier novice teachers not only outperformed their colleagues with regard to student achievement but were also less likely to leave a teaching program prematurely. Altogether, grit seems to play a role in achievement and retention, both of which are highly relevant for success in life. Since the construct grit was introduced, the close relations to the personality trait conscientiousness and to the construct self-control have been addressed on a theoretical and empirical level. Grit correlates moderately with indicators of achievement and is highly correlated with conscientiousness and self-control (Credé, Tyan, & Harms, 2016; Duckworth & Gross, 2014). In the same vein, the metaanalysis by Credé et al. (2016) showed that the predictive power of grit is moderate and that other noncognitive skills even outperform grit in predicting achievements such as GPA (also see Dumfart & Neubauer, 2016).

Next to the question of validity in German-speaking countries, there has been some discussion about the

domain specificity of grit. To date, however, this issue has only seldom been investigated. The Grit-S "attempts to assess behaviors that are reasonably stable across time and situation" implying that "the tendency to pursue long-term goals with passion and perseverance is relatively domain general" (Duckworth & Quinn, 2009, p. 173). It may be reasonably assumed but is yet to be shown that grit can vary between different interests or goals in life, for example, school and hobbies. Bandura (1994) highlighted the fact that personality traits and motivational constructs are often too general to predict specific behaviors sufficiently. In line with this proposition Wigfield (1997) stresses that the domain specificity of constructs marks an important research desideratum. While we are convinced that the validation of the scale is important, the domain-specific approach to grit marks an additional focal point of our study. Duckworth and Quinn (2009) proposed to test this assumption by asking respondents to answer items separately for particular contexts. We dealt with this desideratum by specifically asking for high school students' grit in the context of school as well as their perseverance of effort for the subjects Mathematics and German in addition to domain-general grit. In other words we tested for domain-general grit as in previous studies and introduced a hierarchical approach, a domain-specific (school-specific) grit, and the even more specific perseverance of effort on the school-subject level. Due to the fact that a shift in interests within a given school subject is prevented by the need to follow the curriculum (i.e., topics change too frequently) we only assessed the subscale perseverance of effort for the respective school subjects. Accordingly, with this paper we want to give a first impetus on the possibility to assess school-specific grit and subject-specific perseverance of effort.

Present Research

Fleckenstein, Schmidt, and Möller (2014) translated the original 12-Item Grit Scale into German. The items were translated from English into German and then cross-checked through back translation. This was done by a native English speaker with years of experience working as a qualified editor in educational research. No significant deviations were found between the original and the final translated scale. The resulting scale was termed the BISS-12 Scale (*Beharrlichkeit and Beständiges Interesse* [perseverance and consistency of interest]; the German word *Biss* [English: *bite*] with a symbolic meaning close to *grit*). For this study we used the eight relevant items from the BISS-12 Scale in accordance with the Grit-S by Duckworth and Quinn (2009).

For this paper we first, wanted to investigate the factorial structure and measurement invariance of the German scale as well as the psychometric properties. We assumed a second-order factor structure with the superordinate latent construct grit. We used the data from Study 3 to cross-validate these findings. Our aim was to test the applicability of the German scale in a broader age range. We used samples of adults (Study 1 and Study 2) and a sample of younger high school students (Study 3). The original scale was successfully used in different kinds of samples (e.g., Adults: Von Culin et al., 2014; Adolescents: Robertson-Kraft & Duckworth, 2014; Children: Duckworth et al., 2009).

Second, we used external criteria to investigate the convergent and discriminant validity of the scale. As the theoretical proximity of grit and (1) self-control has been discussed thoroughly (Duckworth & Gross, 2014), we attempted to show the empirical distinctness of the two constructs in latent model tests. The two constructs share similar features and thus can be expected to correlate strongly. However, we still expected the two constructs to be empirically distinct: self-control refers to the inhibition of task-irrelevant impulses in a certain short-term situation, whereas grit means (1) staying interested in superordinate goals over long periods of time and (2) persisting to achieve those higher-order goals even when setbacks occur and alternatives to reach the superordinate goals have to be actively created (for an in-depth discussion of the distinction between grit and self-control, see Duckworth & Gross, 2014; Tangney, Baumeister, & Boone, 2004). Furthermore, we focused on the relationship of grit with other constructs such as (2) the Big Five that have been shown to correlate differentially with grit (Fleckenstein et al., 2014). As stated above, due to the close relatedness of grit and *conscientious*ness significant positive correlations were expected in contrast to the four other personality traits namely neuroticism, extraversion, openness to experience, and agreeableness. Next, since the belief in one's own abilities seems to be an important prerequisite for achieving long-term goals we assumed (3) self-concept to be highly correlated with grit. Another indicator of convergent validity would be a positive relationship between grit and (4) GPA (Duckworth et al., 2009). In light of divergent validity we investigated the relation with school procrastination in Study 3. Students who procrastinate a lot tend to finish their work last minute and sometimes are not able to finish the tasks they aimed to do due to a lack of time (Patzelt & Opitz, 2014). Thus, we expected negative correlations between grit and (5) procrastination.

Finally we wanted to give a first impetus on the research desideratum in respect to the domain specificity of grit as discussed earlier. In this regard, we investigated school-specific grit and subject-specific *perseverance of effort* for the first time.

Study 1

The objective of Study 1 was to examine the factor structure of the BISS-8 Scale (see Table A1 in Appendix for the German version), in accordance with the corresponding findings for the Grit-S, using confirmatory factor analysis (CFA) and to test for measurement invariance for gender. Moreover, we investigated the relationship between grit and self-control in a latent model test.

Materials and Methods

Sample

The sample for Study 1 consisted of N = 525 university students who were enrolled in teacher training programs at several universities across Germany. Their average age was M = 27.93 (SD = 3.63) and 72.1% were female. The questionnaire was administered online as part of the project "Panel study on teacher education" (*Panel zum Lehramtsstudium* [PaLea]; Bauer et al., 2010).

Measures

In accordance with the original version, a 5-point Likert scale was chosen as the response format for the German version of the Grit-S (1 = not at all like me to 5 = very much like me). To assess self-control, we used the German adaptation of the Short Self-Control Scale (Bertrams & Dickhäuser, 2009; Lindner, Nagy, & Retelsdorf, 2015). The scale consists of 13 items (e.g., "*I am in good in resisting temptations*"). Again a 5-point Likert scale with the aforementioned response format was used.

Statistical Analyses

To test the measurement model of the BISS-8 Scale, we applied CFA. Several indices of fit have been suggested to evaluate the goodness of fit for the structural equation model (Marsh, 2007; West, Taylor, & Wu, 2012). For the present analyses, we considered the root mean square error of approximation (RMSEA), the Tucker-Lewis Index (TLI), the Comparative Fit Index (CFI), and the standardized root mean square residual (SRMR). TLI and CFI values greater than .90 or .95 are typically interpreted to reflect an acceptable or excellent fit to the data. RMSEA values lower than .05, .06, or .08 and SRMR values lower than .08 or .10 are typically interpreted to reflect a close or a reasonable fit to the data. All CFA models were estimated using Mplus, Version 7, using the robust full information likelihood estimator (MLR).

Measurement invariance for gender was tested by comparing nested models with different degrees of parameter restrictions (first- and second-order factor loadings, item intercepts, and item residual variances). If the fit of the model with more restrictions (e.g., a model with intercepts,

ſable 1	. Means,	standard	deviations,	item selectivities,	and item	difficulties	for the	BISS-8	Scale in	Study '	1, Study	2, and 3	Study 3	3
---------	----------	----------	-------------	---------------------	----------	--------------	---------	--------	----------	---------	----------	----------	---------	---

		Study 1ª (N = 525 university students)			Study 2ª (N = 173 university students)			Study 3 ^b (N = 271 high school students)					
Item		М	SD	p _{corr}	r _{it}	М	SD	p _{corr}	r _{it}	М	SD	r _{it}	p _{corr}
Consistency of interest													
1.*	l often set a goal but later choose to pursue a different one.	3.71	0.92	.58	.60	3.66	0.84	.56	.38	3.36	1.00	.41	.51
2.*	New ideas and projects sometimes distract me from previous ones.	3.23	1.05	.46	.54	3.08	0.88	.41	.40	3.13	1.04	.39	.46
3.*	I have been obsessed with a certain idea or project for a short time but later lost interest.	3.61	1.13	.57	.46	3.17	1.23	.46	.28	3.45	1.12	.38	.54
4.*	I have difficulty maintaining my focus on projects that take more than a few months to complete.	3.86	1.02	.63	.57	3.61	0.95	.55	.52	3.47	1.11	.42	.55
Perseverance of effort													
5.	I finish whatever I begin.	4.00	0.81	.66	.53	4.03	0.76	.67	.49	3.61	1.02	.62	.58
6.	Setbacks don't discourage me.	3.81	0.89	.61	.34	3.64	0.91	.56	.23	3.66	1.04	.31	.60
7.	l am a hard worker.	3.94	0.91	.65	.54	3.75	0.89	.59	.55	3.00	1.14	.54	.44
8.	I am diligent.	4.00	0.93	.67	.54	3.73	1.01	.59	.53	3.15	1.13	.55	.47

Notes. p_{corr} = Corrected item difficulties; r_{it} = part-whole corrected item discriminations. *Item recoded; ^aResponses were made on a 5-point Likert scale ranging from (1) "not at all like me" to (5) "very much like me". ^bMeans transformed from 4- to 5-point scale to facilitate comparison with means for the samples of Study 1 and Study 2. Items were presented in the German language.

loadings, and residual variances held equal across groups) does not differ substantially from the less restrictive model, a stronger form of invariance can be expected, and serves as an indicator for validity. We followed the ad hoc guide-lines for an evaluation of model fit when testing for measurement invariance presented by Cheung and Rensvold (2002) and Chen (2007). Because the χ^2 -difference test tends to be unreliable in large samples, these authors suggested that support for the more restrictive model requires a change in CFI of less than .01 or a change in RMSEA of less than .015. Finally, it should be pointed out that the TLI and the RMSEA tend to punish a lack of parsimony in a model, that is, a more parsimonious model can reach better fit indices than a less parsimonious model (Marsh, 2007).

Results

Psychometric Properties and Measurement Model

Means, standard deviations, corrected item difficulties, and item selectivities are provided in Table 1. The BISS-8 Scale showed good to excellent corrected item difficulties and item selectivities. The internal consistency was good: Cronbach's $\alpha = .80$.

The CFA for the two-dimensional higher-order model proposed initially by Duckworth and Quinn (2009) did not fit the data sufficiently: $\chi^2(19) = 92.10$, CFI = .92, TLI = .89, RMSEA = .09, SRMR = .08. Due to high modification indices, we allowed the correlation between item



Figure 1. Standardized factor loadings for the second-order model of grit for Study 1/Study 3.

residuals of two items (Item 7 and Item 8). Because of the similarities of the two items, these modifications seemed to be theoretically sound. The modified model appeared to fit the data well: $\chi^2(18) = 25.35$, CFI = .99, TLI = .99, RMSEA = .03, SRMR = .03. All manifest variables and first-order factors loaded significantly (p < .001) and substantially ($\lambda \ge .45$) on the first-order factors and on the second-order factor (see Figure 1).

Measurement Invariance

To test measurement invariance for gender, different levels were compared. Based on the guidelines for model comparisons presented earlier, the assumption of partial strict invariance was supported (see Table 2).

Model	Parameters constrained	χ^2	df	CFI	TLI	RMSEA	SRMR
Measurement inv	ariance in Study 1						
1	None (configural invariance)	44.311	36	.987	.980	.038	.040
2	First-order FL (metric invariance)	54.925	44	.983	.978	.039	.062
3	First- and second-order FL	54.973	45	.984	.981	.037	.062
4	First- and second-order FL, II	78.319	52	.959	.956	.056	.078
4a	First- and second-order FL, partial II	61.306	49	.981	.978	.039	.062
5	First-and second-order FL, partial II, IRV	79.206	57	.965	.966	.049	.134
5a	First- and second-order FL, partial II, partial IRV	63.549	53	.984	.983	.065	.065
Measurement inv	ariance in Study 3						
1	None (configural invariance)	45.577	36	.977	.965	.045	.050
2	First-order FL (metric invariance)	53.793	44	.977	.971	.041	.066
3	First- and second-order FL	53.912	45	.979	.974	.039	.067
4	First- and second-order FL, II	62.349	52	.976	.974	.039	.075
5	First- and second-order FL, II, IRV	69.327	60	.978	.979	.034	.089
5a	First- and second-order FL, II, partial IRV	67.052	59	.981	.982	.032	.76

Table 2. Measurement invariance across gender in Study 1 and Study 3

Notes. FL = factor loadings; II = item intercepts; IRV = item residual variances. For model identification in models 1 and 2 (item intercepts freely estimated) latent means were fixed to zero. CFI = Comparative Fit Index. TLI = Tucker-Lewis index. RMSEA = root mean square error of approximation. SRMR = standardized root mean square residual.

External Criteria

We found strong correlations between grit and selfcontrol, which are presented in Table 3. The correlations between self-control and the subscales *consistency of interest* (r = .61) and *perseverance of effort* (r = .65) did not differ substantially. To investigate whether the two scales still measure two distinct theoretical constructs, we compared a model with the items loading on one common factor with a nested model with the two theoretically postulated factors self-control and the second-order factor grit with its subdomains *consistency of interest* and *perseverance of effort*. The data showed a significantly better fit for the theoretically derived factors model, $\Delta \chi^2 = 288$; df = 5; p < .001.

Discussion

The aim of Study 1 was to establish a structural model using CFA and to assess the psychometric properties of the shortened BISS-8 Scale in accordance with the Grit-S by Duckworth and Quinn (2009). The scale showed good to excellent psychometric properties and the second-order factorial model fit the data well when the correlation between two item residuals was allowed. As expected, the correlation between grit and self-control was strong. However, model comparisons indicated that the scales still assessed two distinct constructs.

Additionally, as an indicator for structural validity, the assumption of partial strict model invariance was supported for gender. Further research is needed concerning the replication of the measurement model, as well as more criteria to assess the external validity. As proposed by Duckworth et al. (2007), the scale should be valid for adults as well as adolescents; therefore, a sample of younger students is necessary.

Study 2

The objective of Study 2 was to investigate the construct and criterion validity of the BISS-8 Scale. We expected it to be strongly related to the personality trait conscientiousness (Roberts, Lejuez, Krueger, Richards, & Hill, 2014). Moreover, grit was expected to be positively correlated with self-efficacy as well as self-concept, since the belief in one's own abilities seems to be an important prerequisite for achieving long-term goals. Grittier students can also be expected to have achieved higher grades in school, which is why we asked them for their GPA.

Materials and Methods

Sample

The sample consisted of N = 173 university students at a university in Northern Germany. Their average age was M = 24.84 (SD = 3.94) and 66.7% were female. The data were collected at the start of lectures on educational psychology, which all students in this university program have to participate in. The administration of the questionnaire took approximately 15 min.

	М	SD	α	r	р
Study 1					
Self-control	3.46	0.63	.87	.73	< .001
Study 2					
Grade point average ^a	2.20	0.49	-	.21	.005
Neuroticism ^b	2.90	0.72	.68	04	.591
Extraversion ^b	3.60	0.84	.83	.09	.262
Openness to experience ^b	3.69	0.75	.74	11	.143
Conscientiousness ^b	3.64	0.67	.74	.70	< .001
Agreeableness ^b	3.21	0.77	.63	.06	.471
Generalized academic self-concept	3.85	0.79	.85	.16	.034
Generalized self-efficacy	3.59	0.43	.81	.31	< .001
Study 3					
Year of schooling	8.15	1.03	-	08	.187
Grade point average ^a	2.61	0.65	.83	.24	< .001
Procrastination	2.22	0.68	.90	47	< .001

Table 3. Internal consistencies of external criteria and manifest correlations with the BISS-8 Scale in Study 1, Study 2, and Study 3

Notes. ^aGrade point average was recoded for better readability (higher values mean better achievements). ^bScales of the BFI-K.

Measures

For this study, grit was assessed using the BISS-8 Scale that we have described in more detail in Study 1. For validation purposes, the questionnaire also included the German short version of the Big Five Inventory (BFI-K; Rammstedt & John, 2005; $\alpha = .63-.83$), assessing the big five: neuroticism (four items; e.g., "I worry a lot"), extraversion (four items; e.g., "I am sociable and can let myself go"), openness to experience (five items; e.g., "I have a wide range of interests"), conscientiousness (four items; e.g., "I complete tasks thoroughly"), and agreeableness (four items; e.g., "I trust people and have faith in the good in people"). In addition, we assessed general self-efficacy (Schwarzer, 1994; $\alpha = .81$) using four items (e.g., "I am not as smart as the others" recoded) and general academic self-concept (Dickhäuser, Schöne, Spinath, & Stiensmeier-Pelster, 2002; $\alpha = .85$) with ten items (e.g., "I think there is a solution to every problem"). A 5-point Likert scale was chosen as the response format for the scales in Study 2 (1 = not at all like me to5 = very much like me). Additionally, participants were asked to provide their gender, and their GPA of the German higher education entrance qualification (Abitur, 1 = very good to 6 = insufficient).

Results

Table 1 shows the psychometric properties for the items of the BISS-8 Scale in our sample. The overall internal consistency was $\alpha = .72$. The subscales were correlated at r = .45. Selectivity for one item was $r_{it} < .30$; however,

removing this item did not lead to a significant increase in Cronbach's α . The mean interitem correlation was r = .42.

The correlations of the BISS-8 Scale with the external validation criteria are shown in Table 3, a full correlation matrix is presented in Table 4. All in all, the directions of the relationships were according to expectations.

The BISS-8 Scale strongly correlated with the Big Five subscale of conscientiousness. The other Big Five dimensions showed no statistically significant correlations with the BISS-8 Scale. In light of the ongoing debate on the predictive power of noncognitive traits mentioned earlier, we regressed GPA on grit. Grit emerged as significant predictor of GPA: $\beta = -.22$, p < .01, B = -.19, SE = .07, $R^2 = .05$. When adding conscientiousness to the regression neither grit nor conscientiousness were significant predictors of GPA: $\beta_{grit} = -.13$, p = .23, $B_{grit} = -.11$, $SE_{grit} = .09$; $\beta_{con} = -.11$, p = .32, $B_{con} = -.08$, $SE_{con} = .08$; $R^2 = .05$. As expected, there was a positive association between grit and general self-efficacy. The correlation with academic self-concept was lower but still significant.

Discussion

The results of Study 2 can be seen as indicators for the reliability and validity of the BISS-8 Scale. The internal consistency and psychometric properties were good; however, there was one item with low selectivity. Since eliminating the item did not improve the internal consistency or the model fit, we decided to keep all eight items in order to

Table 4. Manifest correlations for the scales used in Study 2 and Study 3

Study 2	Ν	E	0	С	А	ASC	SE
1. Grade point average ^a	.16*	.04	.08	.19*	.02	.16*	.02
2. Neuroticism ^b (N)		13	.10	01	04	32**	43**
3. Extraversion ^b (E)			.21**	.12	.11	.14	.39
4. Openness to experience ^b (0)				03	.13	.18*	.15*
5. Conscientiousness ^b (C)					.00	.02	.19*
6. Agreeableness ^b (A)						.01	.03
7. Generalized academic self-concept (ASC)							.37**
8. Generalized self-efficacy (SE)							
Study 3	Gra	ade point ave	rage ^a	Procras	tination		
Year of schooling		.01		.2	8**		
Grade point average ^a				1	6**		

Notes. ^aGrade point average was recoded for better readability (higher values mean better achievements). ^bScales of the BFI-K. *p < .05; **p < .01.

ensure the comparability to the original Grit-S. The correlations with the external validation criteria were mostly according to expectations (gender, GPA, conscientiousness, and generalized self-efficacy).

Study 3

The objective of Study 3 was to replicate the findings regarding the CFA conducted in Study 1 in a younger sample and to get further indicators for the scale's validity. Furthermore, we tested for the domain specificity of grit. In addition, we assessed subject-specific *perseverance of effort*. We expected school-specific grit to correlate higher with GPA than the domain-general measure of grit. Furthermore, subject-specific *perseverance of effort* is expected to be more closely related to the grades in the respective school subjects than domain-general and school-specific grit, respectively.

Moreover, grit was expected to correlate negatively with the students' tendency to procrastinate, as gritty individuals should be inclined to finish their work on time and to not let themselves be distracted from the task at hand by more pleasurable activities.

Materials and Methods

Sample

For Study 3, we collected data from a sample of N = 271 high school students in school year 7 (n = 97; 35.8%), school year 8 (n = 67; 24.7%), school year 9 (n = 76; 28.0%), and school year 10 (n = 31; 11.4%) at an academic track high school (*Gymnasium*) in the federal state of Schleswig-Holstein, Germany. Their average age was M = 13.41 (SD = 1.22) and 55.0% were female. The data were

collected in the schools. The administration of the questionnaire took approximately 20 min.

Measures

In this study, we aimed to replicate the CFA from Study 1 and test measurement invariance for gender for the BISS-8 Scale in a sample of high school students. As in Study 1, we used the robust full information likelihood estimator. All CFA models were estimated using M*plus*, Version 7.

To measure school-specific grit we adapted the BISS-8 Scale by adjusting the preface and the items of the scale to the school context (see the domain-specific and subject-specific scales in Tables A2 and A3 in Appendix). We used subject-specific scales with four items for mathematics ($\alpha = .75$) and German ($\alpha = .79$). For Study 3 a 4-point Likert scale was chosen as the response format (1 = not *at all like me* to 4 = very *much like me*). We adjusted the items of the subject-specific *perseverance of effort* to be subjectspecific, for example, "In mathematics I am diligent."

Finally, we administered a self-developed scale to measure procrastination in school (adapted from the German version of the Aitken Procrastination Scale; Patzelt & Opitz, 2005) using four items per subject (e.g., "I sometimes postpone my homework until time runs out") and calculated a GPA from the students' most recent report cards. The grades have been recoded for better readability.

Results

Means, standard deviations, corrected item difficulties, and item selectivities for the sample of high school students are provided in Table 1. The BISS-8 Scale showed good to excellent corrected item difficulties and item selectivities. The internal consistency was acceptable: Cronbach's $\alpha = .76$.

The two-dimensional higher-order model for the BISS-8 Scale in Study 1 (see Figure 1) was replicated, again allowing the residual correlation between Item 7 and Item 8. The measurement model appeared to fit the data well: $\chi^2(18) = 24.109$, CFI = .988, TLI = .982, RMSEA = .036, SRMR = .032. All items loaded significantly (p < .001)and substantially ($\lambda > .40$) on the latent factor (see Figure 1). Measurement invariance was tested for gender. Based on the guidelines for model comparisons presented in Study 1, the assumption of partial strict model invariance was supported (see Table 2).

The school-specific scale showed good corrected item difficulties and item selectivities. The internal consistency was acceptable: Cronbach's α = .74. As expected the school-specific BISS-8 Scale showed substantial correlations with GPA (r = .31, p < .001). To test for differential effects we predicted GPA by school-specific and domain-general grit applying a multiple regression analysis. Only schoolspecific grit was a significant predictor ($\beta_{school} = .33$, $p < .001, B_{school} = .44, SE_{school} = .13; \beta_{general} = -.19,$ $p = .84, B_{\text{general}} = -.02, SE_{\text{general}} = .12; R^2 = .10$). The same was true for the grade in German ($\beta_{school} = .22, p < .05,$ $B_{\text{school}} = .03, SE_{\text{school}} = .15; \beta_{\text{general}} = -.02, p = .84, B_{\text{general}} =$ -.03, $SE_{general} = .15$; $R^2 = .23$) and mathematics ($\beta_{school} = .33$, $p < .001, B_{school} = .58, SE_{school} = .17; \beta_{general} = -.17, p = .08,$ $B_{\text{general}} = -.29$, $SE_{\text{general}} = .16$; $R^2 = .23$). The correlations between the grades in German and mathematics and the corresponding subject-specific perseverance of effort were significant ($r_{German} = .21, p < .001; r_{math} = .17, p < .001$). To our surprise school-specific grit again explained a significant amount of variance in the grades of the respective school subjects over and above the subject-specific perseverance of effort using multiple regression analyses (German: $\beta_{school} = .17$, p < .05, $B_{school} = .26$, $SE_{school} = .11$; $\beta_{\text{subject}} = .12, p = .10, B_{\text{subject}} = .13, SE_{\text{subject}} = .08; R^2 = .25;$ mathematics: $\beta_{\text{school}} = .26$, p < .001, $B_{\text{school}} = .47$, $SE_{school} = .13; \beta_{subject} = -.12, p = .10, B_{subject} = -.14,$ $SE_{subject} = .12; R^2 = .22).$

The correlations between the BISS-8 Scale and the external criteria are presented in Table 3. All in all, our assumptions were corroborated: We found a strong negative correlation between grit and procrastination.

Discussion

The main aim for Study 3 was to show that the BISS-8 Scale and the corresponding measurement model can also be used to assess younger high school students as well as to test for the possibility to assess domain-specific grit. The psychometric properties of the items and the scale supported this claim. The replication of the measurement model established in Study 1 showed an equally excellent fit for high school students. In particular, the replication of the post hoc modification indicated the structural validity of the scale. Partial strict measurement invariance for gender strengthened the proposed measurement model for grit.

To give a first impetus on the domain specificity of grit we assessed school-specific grit and subject-specific perseverance of effort. In accordance with our expectations, school-specific grit was a more valuable predictor for GPA than domain-general grit as well as for the grades in both mathematics and German. In contrast, the correlations of subject-specific perseverance of effort with the respective school subjects were significant, but regression analyses showed school-specific grit to be a better predictor of the grades in German and mathematics, which was against our expectations. The measure, although pointing in the expected direction, cannot be seen as a comprehensive indicator of subject-specific grit.

The association with procrastination (negative) further supported the validity of the BISS-8 Scale.

General Discussion

The measurement of grit has been examined thoroughly in the US while it is still in very early stages of development in Europe. To our knowledge, validation studies for measures of grit in different languages and contexts in Europe do not exist. Thus, in our studies, we aimed to close this research gap by adapting and closely examining the German version of the most established grit measure: the Grit-S (Duckworth & Quinn, 2009). To take a step further, we adjusted the scale to give a first impetus on the domain specificity of grit.

The aim of the present research was to provide first indicators for the factorial structure, validity, and reliability of the BISS-8 Scale, a German adaptation of the Grit-S. Accordingly, we started off by performing a CFA for the BISS-8 Scale in Study 1. The two-dimensional higher-order measurement model yielded excellent fit indices for the BISS-8 Scale for both university and high school students. It was also shown to be invariant for gender. We were able to replicate these findings in Study 3. All in all, these are strong indicators for the structural validity of the BISS-8 Scale.

When investigating the convergent and discriminant validity of the BISS-8 Scale, we found significant correlations with a number of external criteria: positive correlations with GPA, conscientiousness, self-efficacy, and self-control as well as a negative correlation with procrastination supported the scale's validity. The fact that, although strongly correlated, grit still seems to differ from the similar construct of self-control also empirically reinforces its relevance as a disjunct psychological construct.

Presenting findings on the domain specificity of grit for the very first time we were able to show that school-specific grit is more closely related to GPA than domain-general grit. Thereby, we were able to give a promising first impetus regarding the assumption of Duckworth and Quinn (2009) and the desideratum pointed out by Bandura (1994) and Wigfield (1997) in regard to the domain specificity of constructs as an important research desideratum. Additionally these findings can help to shed light on the lacking incremental validity of domain-general grit over and above conscientiousness - which we were able to corroborate with our findings from Study 2 - as mentioned by Ivcevic and Brackett (2014) and the meta-analysis by Credé et al. (2016). Both assumed grit to be a better predictor of more narrow goals. All in all, our results show that it is worthwhile to assess domain-specific grit in a school context. Further research concerning other domains could underpin our findings.

Limitations and Future Directions

Finally, some limitations of the present study need to be addressed. First, our samples only comprised students from academic track high schools and universities. High means and somewhat restricted variance may have reduced the size of some correlations between the BISS-8 Scale and the external criteria. This suggests that our findings may be a rather conservative approximation of the true correlations. Second, it is not possible to further evaluate the generalizability of the findings at this point due to the positively selected sample. Third, longitudinal surveys are needed to provide indicators for, e.g. predictive validity. Moreover, as was mentioned for the original grit scale, we only used self-report measures to assess grit, the limits of which are well documented (Lucas & Baird, 2006). In addition, Duckworth and Yeager (2015) pointed out that great care has to be taken, when measuring traits like grit. Not only because of the quality of the self-report measures but also because of the impact these concepts can have when used improperly.

As already mentioned, we did not assess subject-specific grit comprehensively, but only the *perseverance of effort* for the respective school subjects. We only can assume this to be the reason for it to fall behind school-specific grit in predicting the respective grades. To measure the *consistency of interest* in a given school subject is difficult because curricula do not allow a switch of interests as it would be possible regarding hobbies or even the workplace. To measure grit in an educational setting on an even lower hierarchical level (i.e., subject-specific grit) presents a challenge for further research. In addition, we were not able to assess the stability of grit on a lower hierarchical level. Research tapping this desideratum still needs to be conducted.

Ivcevic and Brackett (2014) argued that grit would be best suited for easy-to-define, narrow long-term goals that require commitment and focused practice. They recommend the broader trait conscientiousness to predict broader success criteria like GPA. Unfortunately, our data did not allow analyses regarding the incremental predictive value of domain-general and school-specific grit over and above conscientiousness. Nevertheless, in Study 2 we were able to show that neither grit nor conscientiousness were significant predictors of GPA in a multiple regression analysis. In this regard, it should be mentioned that the meta-analysis by Credé et al. (2016) showed that grit cannot be expected to be a construct independent of conscientiousness. Even though we have to highlight the fact that the measure used to assess conscientiousness can have great impact on the results, studies investigating the claim of a relationship on a subdomain level still remain to be conducted. Nevertheless, Credé et al. (2016) argue that the construct may be useful beyond conscientiousness in settings where retention is uncertain.

The measurement model of the BISS-8 Scale posed a challenge: A correlation between two items residuals ("I am a hard worker" and "I am diligent") had to be allowed in order to improve the model fit in Study 1. We were able to replicate this modification in Study 3; this can be seen as an indicator for the model's validity. Duckworth and Quinn (2009) did not report a similar problem; this might be due to semantic differences between the term *diligent* and its German translation *fleißig*. While diligent is defined as "steady, earnest, and energetic effort" ("diligent," 2015), the meaning of *fleißig* is narrower and may be more appropriately translated as "industrious" or "hard-working" ("fleißig," 2015).

An issue that pertains to the BISS-8 Scale as well as to its original version is that one has to consider the possibility that the two first-order factors identified were an artifact of positively and negatively scored items. This was already pointed out by Duckworth et al. (2007), who were nevertheless convinced that this factor structure reflects two conceptually distinct dimensions. However, further research with only positively coded items in both subscales could support this claim empirically.

Acknowledgments

Parts of the research reported in this article are part of the project "Panel study on teacher education" (Panel zum Lehramtsstudium [PaLea]). The project is funded by the Federal Ministry of Education and Research (BMBF; promotional reference: 01PH08004 and 901PH13001A).

References

- Akın, A., & Arslan, S. (2014). The relationships between achievement goal orientations and grit. *Education and Science*, 39, 267–274. doi: 10.15390/eb.2014.2125
- Arslan, S., Akın, A., & Çitemel, N. (2013). The predictive role of grit on metacognition in Turkish university students. *Studia Psychology*, 55, 311–320.
- Bandura, A. (1994). Regulative function of perceived self-efficacy. In M. G. Rumsey, C. B. Walker, & J. H. Harris (Eds.), *Personal* selection and classification (pp. 261–271). Hillsdale, NJ: Erlbaum.
- Bauer, J., Drechsel, B., Retelsdorf, J., Kauper [Sporer], T., Rösler, L., Prenzel, M., & Möller, J. (2010). Panel zum Lehramtsstudium – PaLea: Entwicklungsverläufe zukünftiger Lehrkräfte im Kontext der Reform der Lehrerbildung [Panel study on teacher education – PaLea. Professional development of prospective teachers under the conditions of the bologna process]. Beiträge zur Hochschulforschung, 32, 34–55.
- Bertrams, A., & Dickhäuser, O. (2009). Messung dispositioneller Selbstkontroll-Kapazität. Eine deutsche Adaptation der Kurzform der Self-Control Scale (SCS-K-D) [Measuring dispositional self-control capacity. A German adaptation of the short form of the Self-Control Scale (SCS-K-D)]. Diagnostica, 55, 2–10. doi: 10.1026/0012-1924.55.1.2
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14,* 464–504. doi: 10.1080/ 10705510701301834
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-offit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 233–255. doi: 10.1207/S15328007SEM0902_5
- Cox, C. M. (1926). Genetic studies of genius: Vol. 2. The early mental traits of three hundred geniuses. Stanford, CA: Stanford University Press.
- Credé, M., Tynan, M. C., & Harms, P. D. (2016). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*. Advance online publication. doi: 10.1037/pspp0000102
- Dickhäuser, O., Schöne, C., Spinath, B., & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept. Konstruktion und Überprüfung eines neuen Instrumentes [Skalen zum akademischen Selbstkonzept (SASK). Construction and evaluation of a new tool assessing the academic selfconcept]. Zeitschrift für Differentielle und Diagnostische Psychologie, 23, 393–405. doi: 10.1024/0170-1789.23.4.393
- Diligent. (2015). In Merriam-Webster.com. Retrieved from http:// www.merriam-webster.com/dictionary/diligent
- Duckworth, A., & Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science*, 23, 319–325. doi: 10.1177/0963721414541462
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92, 1087–1101. doi: 10.1037/0022-3514.92.6.1087
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). Journal of Personality Assessment, 91, 166–174. doi: 10.1080/00223890802634290
- Duckworth, A. L., Quinn, P. D., & Seligman, M. E. P. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology*, 4, 540–547. doi: 10.1080/17439760903157232
- Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98, 198–208. doi: 10.1037/0022-0663.98.1.198

- Duckworth, A. L., Tsukayama, E., & Geier, A. B. (2010). Selfcontrolled children stay leaner in the transition to adolescence. *Appetite*, 54, 304–308. doi: 10.1016/j.appet.2009.11.016
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44, 237–251. doi: 10.3102/0013189X15584327
- Dumfart, B., & Neubauer, A. C. (2016). Conscientiousness is the most powerful noncognitive predictor of school achievement in adolescents. *Journal of Individual Differences*, 37, 8–15. doi: 10.1027/1614-0001/a000182
- Fleckenstein, J., Schmidt, F. T. C., & Möller, J. (2014). Who's got grit? Perseverance and consistency of interest in preservice teachers. A German adaptation of the 12-Item Grit Scale. *Psychologie in Erziehung und Unterricht*, 61, 281–286. doi: 10.2378/peu2014.art22d
- fleißig. (2015). *In Collins dictionaries*. Retrieved September 28, 2015, from http://www.collinsdictionary.com/dictionary/german-english/fleißig
- Galton, F. (1892). Hereditary Genius: An inquiry into its laws and consequences. London, UK: Macmillan.
- Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality*, 52, 29–36. doi: 10.1016/j.jrp.2014.06.005
- Lindner, C., Nagy, G., & Retelsdorf, J. (2015). The dimensionality of the Brief Self-Control Scale-An evaluation of unidimensional and multidimensional applications. *Personality and Individual Differences*, 86, 465–473. doi: 10.1016/j.paid.2015.07.006
- Lucas, R. E., & Baird, B. M. (2006). Global self-assessment. In M. Eid & E. Diener (Eds.), Handbook of multimethod measurement in psychology (pp. 29–42). Washington, DC: APA. doi: 10.1037/11383-003
- Marsh, H. W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *Journal of Educational Psychology*, 99, 775–790. doi: 10.1037/0022-0663.99.4.775
- Patzelt, J., & Opitz, I. (2005). Deutsche Version der Aitken Procrastination Scale (APS-d) [Aitken Procrastination Scale (APS; Aitken, 1982) – German version]. In A. Glöckner-Rist (Ed.), ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente (8th ed.). Mannheim, Germany: ZUMA.
- Patzelt, J., & Opitz, I. (2014). Deutsche Version der Aitken Procrastination Scale (APS-d) [German version of the Aitken procrastination scale (APS-d)]. In Zusammenstellung sozialwissenschaftlicher Items und Skalen. Leipzig, Germany: GESIS. doi: 10.6102/zis111
- Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit [Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality]. *Diagnostica*, 51, 195–206. doi: 10.1026/0012-1924.51.4.195
- Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is conscientiousness and how can it be assessed? *Developmental Psychology*, 50, 1315–1330. doi: 10.1037/a0031109
- Robertson-Kraft, C., & Duckworth, A. L. (2014). True grit: Traitlevel perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. *Teachers College Record* (1970), 116, 1–27.
- Schwarzer, R. (1994). Optimistische Kompetenzerwartung: Zur Erfassung einer personalen Bewältigungsressource [Generalized self-efficacy: Assessment of a personal coping resource]. *Diagnostica*, 40, 105–123.

- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High selfcontrol predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271–324. doi: 10.1111/j.0022-3506.2004.00263.x
- Terman, L. M., & Oden, M. H. (1947). The gifted child grows up: Twenty-five years' follow-up of a superior group. Oxford, UK: Stanford University Press.
- Von Culin, K. R., Tsukayama, E., & Duckworth, A. L. (2014). Unpacking grit: Motivational correlates of perseverance and passion for long-term goals. *The Journal of Positive Psychology*, 9(4), 306–312. doi: 10.1080/17439760.2014.898320
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: Guilford Press.
- Wigfield, A. (1997). Reading motivation. A domain-specific approach to motivation. *Educational Psychologist*, 32(2), 59–68. doi: 10.1207/s15326985ep3202_1

Zimmerman, B. J., & Schunk, D. H. (2011). Handbook of selfregulation of learning and performance Educational psychology handbook series. New York, NY: Routledge.

Received December 3, 2015 Revision received October 11, 2016 Accepted October 17, 2016 Published online June 28, 2017

Fabian T. C. Schmidt

IPN - Leibniz Institute for Science and Mathematics Education Olshausenstr. 6224118 Kiel Germany fschmidt@ipn.uni-kiel.de

Appendix

Table A1. BISS-8 Skala	
Beständiges Interesse	
1.	Ich setze mir oft ein Ziel, entscheide mich dann aber später doch ein anders Ziel zu verfolgen.
2.	Neue Ideen und Projekte halten mich manchmal von vorherigen ab.
3.	Ich war schon einmal für eine kurze Zeit von einem Projekt oder einer Idee besessen, habe später aber das Interesse verloren.
4.	Ich habe Schwierigkeiten auf Projekte fokussiert zu bleiben, wenn diese mehrere Monate dauern.
Beharrlichkeit	
5.	Alles was ich beginne, bringe ich auch zu Ende.
6.	Von Rückschlägen lasse ich mich nicht entmutigen.
7.	Ich bin ein hart arbeitender Mensch.
8.	Ich bin fleißig.

Anmerkungen: Items werden anhand einer fünfstufigen Likertskala erhoben (1 = trifft überhaupt nicht zu, 2 = trifft eher nicht zu, 3 = teils teils, 4 = trifft eher zu, 5 = trifft völlig zu). Items der Subskala Beständiges Interesse sind zu rekodieren.

Table A2. School-specific grit items in English and German

	Items in English language	Items in German language				
	Consistency of interest	Beständiges Interesse				
1.	l often set a goal in school but later choose to pursue a different one.	In der Schule nehme ich mir oft etwas vor, verliere mein Ziel dann aber aus den Augen.				
2.	New ideas and plans sometimes distract me from my goals in school.	Neue Ideen und Vorhaben halten mich manchmal davon ab, meine schulischen Ziele zu verwirklichen.				
3.	In school I have been obsessed with a certain plan or project for a short time but later lost interest.	lch war in der Schule schon mal für eine kurze Zeit von einem Vorhaben oder einem Projekt begeistert und habe dann aber später das Interesse verloren.				
4.	I have difficulty maintaining my focus on school projects that take more than a few weeks to complete.	Wenn schulische Vorhaben oder Projekte länger als einige Wochen dauern, habe ich Schwierigkeiten mich durchgängig darauf zu fokussieren.				
	Persistence of effort	Beharrlichkeit				
5.	In school I finish whatever I begin.	Was ich mir in der Schule vorgenommen habe, mache ich auch zu Ende.				
6.	In school setbacks don't discourage me.	Von Rückschlägen in der Schule lasse ich mich nicht entmutigen.				
7.	I work hard for school.	Ich arbeite hart für die Schule.				
8.	I am a diligent student.	Ich bin ein fleißiger Schüler/eine fleißige Schülerin.				

	Items in English language	Items in German language
	Persistence of effort in mathematics	Beharrlichkeit im Fach Mathematik
1.	In mathematics I finish whatever I begin.	Wenn ich in Mathematik mit etwas beginne, mache ich das auch zu Ende.
2.	In mathematics setbacks don't discourage me.	Von Rückschlägen in Mathematik lasse ich mich nicht entmutigen.
3.	In mathematics I work hard.	In Mathematik arbeite ich hart.
4.	In mathematics I am diligent.	In Mathematik bin ich fleißig.
	Persistence of effort in German	Beharrlichkeit im Fach Deutsch
1.	In German I finish whatever I begin.	Wenn ich in Deutsch mit etwas beginne, mache ich das auch zu Ende.
2.	In German setbacks don't discourage me.	Von Rückschlägen in Deutsch lasse ich mich nicht entmutigen.
3.	In German I work hard.	In Deutsch arbeite ich hart.
4.	In German I am diligent.	In Deutsch bin ich fleißig.