How valid is grit in the postsecondary context?

A construct and concurrent validity analysis

Kevin Fosnacht

Keeley Copridge

Shimon A. Sarraf

Indiana University Center for Postsecondary Research

This is a post-peer-review, pre-copyedit version of an article published in *Research in Higher Education*. The final authenticated version is available online at:

<https://doi.org/10.1007/s11162-018-9524-0>

Corresponding author:

Kevin Fosnacht, Ph.D.

Indiana University Center for Postsecondary Research

1900 East Tenth Street, Suite 419

Bloomington, IN 47406-7512

812-856-1961

kfosnach@indiana.edu

How valid is grit in the postsecondary context? A construct and concurrent validity analysis

Abstract

College admissions leaders increasingly desire to incorporate non-cognitive factors like grit into admissions decisions. Consequently, we examined the validity of the Short Grit Scale (Grit-S) using data collected on undergraduates attending 38 colleges. Using confirmatory factor analysis (CFA), we found that Grit-S does not possess adequate model fit; however, a modified version of the scale does possess this property. Using multi-group CFA, we also found that Grit-S is relatively invariant across multiple demographic groups. Next, we examined the concurrent validity of grit with students' engagement, perceived gains, time spent studying, and grades. These results confirmed previous research that most of the predictive power of grit is contained in its perseverance of effort, not consistency of interest, dimension.

Keywords: Grit; non-cognitive traits; validity; higher education; admissions; measurement invariance

How valid is grit in the postsecondary context? A construct and concurrent validity analysis

In the search to build a successful student body, admissions leaders at postsecondary institutions have focused on identifying skills and traits beyond grades and standardized test scores to recognize students with the potential to succeed at their institutions. One non-cognitive factor with this potential is grit. The increasing popularity of grit, coined by Angela Duckworth, has encouraged admissions officers to review non-cognitive factors to build a diverse class and create an engaging campus community (Akos and Kretchmar 2017; Powell 2013; Sedlacek 2017; Wick 2015). Grit has the potential to assist colleges in selecting a better-rounded student body as it may help identify students from historically marginalized backgrounds with a substantial likelihood of success. Due to grit's focus on stamina and commitment to long-term goals, characteristics pertinent to college retention and completion, grit can theoretically identify students with a high probability of succeeding despite some deficiency in the criteria traditionally used in admissions decisions and help diversify the student body.

The focus on grit has expanded beyond the admissions and research communities. The United States Department of Education promoted using grit as a tool to prepare future generations for college and beyond (US Department of Education 2013). Furthermore, grit has been the focus of a TED talk with over 12 million views (Duckworth 2013), numerous articles in the popular press (e.g., Del Giudice 2014; Engber 2016), and earned Duckworth the MacArthur "Genius" grant (MacArthur Foundation 2017). Despite the growth in the popularization of grit, the concept has not been exposed to the empirical scrutiny it deserves if it is going to be a factor in high stakes decisions within the educational research community. Duckworth, Peterson, Matthews, and Kelly's (2007) initial research on grit has examined this concept with high-

achieving populations such as West Point cadets and Scripps Spelling Bee finalists, populations that are not representative of the typical student. Furthermore, much of the current research examines smaller populations ranging in the hundreds. Consequently, in this study, we utilized the National Survey of Student Engagement's ability to collect data on over 11,000 undergraduates to test the construct validity of grit and its concurrent validity for measures of engagement, self-perceived gains, time use, and GPA using a sample that represents a more typical college student population.

Literature Review

Grit is the "perseverance and passion for long-term goals" (Duckworth et al. 2007, p. 1087). Grit is a non-cognitive personality trait that is operationalized as a high-order construct with two lower order features, perseverance of effort and consistency of interest (Duckworth et al. 2007; Duckworth and Quinn 2009). Perseverance of effort refers to an individual's tendency to work hard in the face of setbacks or obstacles while the latter, consistency of interest, is the tendency not to change goals and interest frequently (Duckworth et al. 2007; Duckworth and Quinn 2009). These facets theoretically work together to influence an individual's attitude and behavior towards long-term goals. One of the attractive features of grit is the lack of correlation with other measures of intelligence and that it is a trait that can be potentially changed (Duckworth 2016; Duckworth et al. 2007). Consequently, grit can theoretically help all students succeed. Previous research has correlated grit with outcomes like persistence in higher education and success in long-term, difficult tasks (Duckworth et al. 2007, Duckworth and Quinn 2009). However, the grit concept has been criticized from a variety of perspectives (Credé, Tynan, and Harms 2016; Ris 2015). Below, we summarize grit and its two sub-constructs, perseverance of

effort and consistency of interest, review previous research of how grit influences outcomes, and discuss multiple criticisms of grit.

Grit and Academics

Studies focusing on grit among college students show that grittier students more frequently persist and succeed academically (Bowman, Hill, Denson, and Bronkema 2015; Cross 2014; Duckworth 2007; Strayhorn 2014). Duckworth and colleagues' (2007) study on grit and various high-achieving populations such as the West Point cadets in training and Scripps Spelling Bee finalist found that grit predicted completion of their training program for the cadets and the number of hours the spelling bee students practiced after holding constant other factors. Bowman et al. (2015) found that grit was positively correlated with academic adjustment, GPA, satisfaction, sense of belonging, and more frequent faculty-student interaction and co-curricular engagement. However, they found that these effects were attributable to the perseverance of effort dimension, not consistency of interest. Strayhorn's (2014) study examined grit using a sample of African American males and found that grittier black males had higher college grade point average than their counterparts. Cross' (2014) study examined the relationship between grit scores, GPA, and gender. The results showed that there was a significant and positive relationship between grit and GPA for females but not for males. He also discovered a positive relationship between grit and doctoral degree attainment (Cross 2014). More recently, Akos and Kretchmar (2017) found that grit was correlated with first-year GPA, but also that the relationship was primarily attributable to the perseverance of effort dimension. Also, they found a negative relationship with changing majors, no relationship to credit accumulation and, using a family and peer assessment of student grit, a positive relationship to first-year GPA and credit accumulation.

Wolter and Hussain (2014) investigated grit and its relations to college students' self-regulated learning and academic achievement. They defined self-regulated learning as the process in which "students take an active, purposeful role in managing motivational, cognitive, and behavioral aspects of their learning" (Wolter and Hussain 2014, p. 295). This management consisted of students engaging in different sub-processes such as goal setting, the activation of prior knowledge, progress monitoring, engagement and regulation of learning strategies, and reflection to learn. Their results showed that grittier students were less likely to procrastinate and had reduced levels of delay in the beginning and completing academic tasks. They concluded self-regulated learning appears to mediate the relationship between grit and academic performance.

In contrast, other studies have not demonstrated a correlation between grit and academic outcomes (Bazelaïs et al. 2016; Cross 2013; Stewart 2015). Stewart (2015) found that high school GPA and test scores were predictors of college academic performance but not grit. This finding was replicated by Bazelaïs, Lemay, and Doleck (2016) who examined the relationship between grit and GPA and final exam scores for students taking a gateway physics course. Among graduate students, Cross (2013) found no relationship between grit and doctoral students' dissertation completion.

Recent studies on grit have examined its cross-cultural applicability, an essential facet if admissions staff seek to use grit as a way to boost enrollment of students from historically marginalized communities. Datu, Valdez, and King (2015) examined grit's validity for a sample of Filipino undergraduate and high school students from a collectivist culture. Their results showed that grittier students were more likely to achieve higher satisfaction in life and emotional well-being. Additionally, these effects comport with Bowman and colleagues' results in that the

effects of grit appear to be isolated to the perseverance of effort dimension. O’Neal and colleagues (2014) examined the relationships between grit, stress, depression, and GPA using a sample of documented and undocumented Latinx first-generation students. They found high levels of grit among both documented and undocumented Latinx students. Additionally, the relationship between grit and depression was negative as undocumented Latinx students with lower levels of grit were more likely to report higher levels of depression than documented Latinx students.

Grit and Engagement

To date, limited research has focused on grit’s influence on engagement in educationally purposeful activities (Hodge et al. in press; Robinson 2015). As mentioned above, Bowman and colleagues (2015) found that grit’s perseverance of effort dimension was positively correlated with both faculty-student interaction and co-curricular engagement. Holbein and colleagues (2016) studied the relationship between grit and middle and high school students’ school and civic engagement. Their findings show that grittier students performed noticeably better on a standardized test, had higher levels of school attendance, and a stronger belief in their future engagement in the political process. Hodge and colleagues (in press) study about the engagement of Australian university students found a positive relationship between grit and engagement, which positively correlated with better academic productivity. Moreover, Robinson (2015) found a strong association between grit and engagement in coursework for nursing students.

Critiques of Grit

Both the research literature and popular press has hailed grit as the solution to numerous problems in education (e.g., Duckworth et al. 2007; Duckworth et al. 2009; Del Giudice 2014; Engber 2016; Powell 2013; Sedlacek 2017; Wick 2015). However, these claims about grit have

not always stood up to scrutiny. Credé and colleagues (2016) conducted a meta-analysis of 88 studies focusing on grit and critiqued grit's validity. They claim that the relationship between grit and success has been overstated by Duckworth, as the correlation of grit with academic success is less than .20 in their meta-analysis.

Credé and colleagues (2016) also contest the uniqueness of grit due to its correlation (.84) with the conscientiousness dimension of the Big Five personality traits and suggest that grit is an old, but repackaged concept. Individuals that are conscientious are “thorough, careful, reliable, organized, industrious and self-controlled” (Duckworth et al. 2007, p. 1089). Individuals that are conscientiousness are highly achievement-oriented, a hallmark of grit (Cross 2014). However, grit's proponents argue that grit includes self-control traits that makeup conscientiousness and also focuses on long-term stamina rather than short-term intensity (Cross 2014; Duckworth et al. 2007).

Others have contested the construct validity of grit. Credé and colleagues (2016) point out that Duckworth et al.'s confirmatory factor analysis (CFA) of the Short Grit Scale was invalid. Their CFA model used a higher-order factor structure with two first-order factors which produce an unidentified model without the imposition of unusual constraints¹. Muenks, Wigfield, Yang, and O'Neal (2017) further examined the factor structure of grit, its relation to other constructs, and the best factor structure model of grit. In their study, high school and college students' grit, conscientiousness, self-control, cognitive regulations, effort regulation, and behavioral engagement and disaffection were measured via a survey. Results showed that for

¹ Higher-order factors require three subscales for a CFA model to be identified, without the imposition of unusual constraints. When Duckworth and Quinn's (2009) model is replicated in AMOS 24 (Duckworth and Quinn used version 6.0), the following error message is produced along with the results: “Minimization was unsuccessful. The results that follow are therefore incorrect. The model is probably unidentified. In order to achieve identifiability, it will probably be necessary to impose 2 additional constraints.” Consequently, the results associated with the published higher-order factor grit solution are highly suspect, especially given that a standardized loading is greater than 1.

high school students the two correlated-factor model was the best fit for measuring grit while the bi-factor model was the best fit for college students. Consequently, their findings indicate that the grit scale is configurally variant between different populations condition, suggesting that Duckworth's grit scales lack construct validity. Moreover, Muenks and colleagues (2017) found that grit and its subscales significantly overlapped with personality, self-regulation and engagement literature.

Grit has alternately been critiqued from a class-reproduction standpoint as it could help legitimize existing inequalities throughout society (Gonzales-Stokas 2015; Ris 2015; Socal 2014). This legitimization of existing inequalities can be reproduced through the fundamental attribution error or the tendency to overvalue personality-based explanations for behaviors and situations while ignoring the institutional and systemic constructs that act as barriers to an individual's aspirations (Gonzalez-Stokas 2015). An example of fundamental attribution error in the education system is the unethical action of telling "children who face a society of entrenched economic inequality, that achievement is the result of individual effort and is disconnected from systemic privilege" (Gonzalez-Stokas 2015, p. 516).

Despite these critiques, grit has captured the imagination of college administrators, policymakers, and the popular press. Consequently, it is essential to investigate the efficacy of grit's relationship with college outcomes. The study of grit has examined various types of populations and circumstances, yet these analyses have primarily focused on a narrow population of interest (i.e., high achieving students at a military academy or African American males) or students attending a single institution. Additionally, there is limited research about grit's influence on undergraduate student engagement (Wang and Degol 2016). This lack of research is concerning, as it limits the potential to understand how grit influences student outcomes. As grit

is not correlated with intelligence, grit must be related to behavioral outcomes if it ultimately influences academic success. Consequently, we chose to fill in these research gaps by investigating the construct validity of the most popular grit scale version and its concurrent validity by investigating its association with students' engagement in effective educational practices, perceived gains, time usage, and GPA for a diverse, multi-institutional sample of undergraduates attending bachelor's degree-granting institutions.

Conceptual Framework

At its core, this study is an investigation into the validity of grit. Our validity inquiries were guided by Messick's (1989) unified validity framework. In the framework, Messick moved beyond Cureton (1951) and Cronbach's (1971) notions of validity which focus on the characteristics of a measure. While traditional criterion validity is an essential aspect of Messick's framework, he also emphasized the need to focus on how a measure is interpreted and used in practice. Messick (1995) offers a progressive matrix of validity that ranges from (1) construct validity, (2) construct validity and relevance/utility, (3) construct validity and value implications, (4) construct validity, relevance/utility, value implications, and social consequences. Messick's notion of validity has gone on to inform the current *Standards for Educational and Psychological Testing* (American Educational Research Association et al. 2014).

From this unified view of validity, it is vital to assess multiple facets of validity in relationship to grit. The questions to inquire about include:

- Whether the Short Grit Scale measures grit (the latent construct)?
- Is the theoretical basis of grit sound?
- Does grit have concurrent, discriminant, and predictive validity?

- Is grit generalizable across groups? Does grit discriminate against certain populations?
- What are the risks associated with using grit in a practical setting?

Until these questions are fully answered, using grit in high stakes decisions may lead to unintended consequences.

Purpose

In this study, we sought to examine the construct and concurrent validity of the Short Grit Scale for a large, diverse, multi-institutional sample of college students. To analyze the construct validity of grit, we utilized confirmatory factor analysis (CFA) to test if the scale meets generally accepted thresholds for construct validity using the largest and most diverse data set on grit collected (Credé et al. 2017). Additionally, we conducted multi-group CFAs to test if the relationships differ between subpopulations of students to investigate the measurement invariance of the Short Grit Scale. Next, we examine the concurrent validity of the grit scale by examining its relationship with process indicators of student engagement, self-perceived gains, time spent studying, and GPA. As the hallmarks of grit are perseverance of effort and consistency of interest, we hypothesize that grit is a factor leading to student effort and time dedicated to studying, which ultimately leads to learning and development (McCormick et al. 2013). However, to date, this relationship has not been extensively tested. Consequently, we examined the association between the psychological concept of grit and the behavioral dimensions of student engagement in effective educational practices. Establishing this link between the psychological realm and actual behaviors is important to understanding how students learn. Additionally, if this association does not exist, it calls into question the validity of grit as a standalone concept.

Methods

Data

We utilized data from the National Survey of Student Engagement (NSSE) administered in the winter and spring of 2016. NSSE is a multi-institutional study of first-year and senior students attending bachelor's-granting institutions that examines how often students engage in educationally beneficial activities, students' time-use, and their perceptions of the campus environment. Due to our focus on grit, we focused our analyses on students who attended one of 38 institutions that received a supplemental set of items that included the Short Grit Scale (Duckworth and Quinn 2009). These institutions were randomly selected from a pool of institutions that did not elect to append two additional item sets (topical modules and consortium items) to the core NSSE instrument. A total of 4,668 first-year and 7,082 senior students responded to the item set. The response rate for first-year and senior students was 21% and 24%, respectively. Previous research has demonstrated that NSSE data is reliable at these response rates (Fosnacht, Sarraf, Howe, and Peck 2017).

The students in our sample attended a diverse set of institutions. Roughly a quarter attended institutions that awarded doctoral degrees, over half attended master's colleges and universities, 15% were enrolled at baccalaureate colleges, and 5% attended special focus institutions. Just over 40% of the sample was enrolled at a public college. A plurality of students attended institutions with undergraduate enrollments between 5,000 and 9,999 students. Half of the students attended institutions with a Barron's rating of competitive. A third of the sample attended a minority-serving institution.

Slightly less than two-thirds of the sample were female. White students comprised 57% of the respondents. Asian, Black, Latinx, and multi-racial students represented 6%, 11%, 12%, and 7% of the sample respectively. Five percent of the sample was international. Slightly over a

quarter of the sample did not have a parent who enrolled in college. Twelve percent of the students had parents with some college education. One in ten students had a parent who received an associate's degree. Over a quarter of the sample had a parent who earned a bachelor's degree. Slightly less than one in four students had a parent with a master's degree or higher.

We utilized two sets of key variables in our analyses. First, grit was represented by the eight items in Duckworth and Quinn's (2009) Short Grit Scale. These items were lightly edited from the original scale to conform to NSSE's style and appended to the end of NSSE (see Appendix A for the items wording). The changes were primarily in the response options from very much, mostly, somewhat, not much, and not "like me at all" to NSSE's standard options of not at all, not much, somewhat, mostly, and very much. This change ensured that the options comported with the options on the core NSSE instrument and reduced the cognitive burden on the respondent. We created subscale scores by taking the means of the component items for the items included in the consistency of interest and perseverance of effort subscales. Additionally, we reverse coded the items in the consistency of interest subscale, so that they indicate higher levels of this trait.

Our second set of key variables were process indicators of student learning and development en route to students' ultimate goal of completing college. These variables were represented by 9 of the 10 NSSE Engagement Indicators (we excluded Effective Teaching Practices, as this scale focuses on instructors' efforts), a perceived gains scale, time spent preparing for class, and self-reported grades. We chose to focus on measures of engagement as they have been found to predict students' learning and development and retention (Kuh et al. 2006; National Survey of Student Engagement, n.d.; Pascarella et al. 2010). Additionally, students' grades have long been a proxy for student learning and development. Information on

reliability and validity of the Engagement Indicators is available in NSSE's (2017) Psychometric Portfolio. The items in the perceived gains scale asked about how much their college experience contributed to students' knowledge, skills, and personal development in 10 areas. The Cronbach's α for the perceived gains scale was .91 for first-year students and .90 for seniors. Students' time spent preparing for class was captured in ranges (0, 1-5, 6-10, 11-15, 16-20, 21-25, 26-30, more than 30 hours per week) and recoded to the midpoint (the top category was set to 33 hours per week). Students were asked to report their typical grades in the following categories: A, A-, B+, B, B-, C+, C, C- or lower. We recoded these values to reflect the typical 4.0 GPA scale. To aid in the interpretation of the results, we standardized both grit subscales, the Engagement Indicators, perceived gains scale, time spent preparing for class, and grades to have a mean of 0 and standard deviation of 1.

In addition to these key variables, we also used data on a variety of control variables: race/ethnicity, sex, standardized test score (SAT/ACT), parental education, academic major, greek-life participation, age, athletics participation, transfer status, part-time status, educational aspirations, and on-campus residency.

Analyses

Construct Validity. We began our analyses by assessing the construct validity of the Short Grit Scale using confirmatory factor analysis. Duckworth and Quinn (2009) propose the Short Grit Scale as a higher order factor with two subscales: consistency of interest and perseverance of effort. However, this model is unidentified without the imposition of multiple unusual constraints (see the critiques of grit section for more details). Therefore, we conducted a CFA using MPLUS where the subscales were correlated and not subsumed under a higher-order factor. The model was estimated using full information maximum likelihood. We assessed model

fit using the following standards (Hu and Bentler 1999). The model was judged to have good fit if the comparative fit index (CFI) and Tucker-Lewis index (TLI) were greater than or equal to .95 and the root mean square error of approximation (RMSEA) was less than or equal to .05. The thresholds for adequate fit used were CFI and TLI were greater than or equal to .90, and the RMSEA was less than or equal to .06. Also, we report the χ^2 results for the models; however, due to our large sample size and the sensitivity of the χ^2 statistic to sample sizes, we focused our interpretation on the CFI, TLI, and RMSEA.

Next, we examined the measurement invariance of the Short Grit Scale by replicating the confirmatory factor analysis by subgroup. Confirmation of measurement invariance helps ensure valid inferences from between-group comparisons (Horn & McArdle, 1992). If grit has applicability for high-stakes decisions, like admissions, it is critical to assess whether the scale's measurement properties vary by subgroups, as the use of grit in such decisions could unfairly advantage or disadvantage particular groups of students. Measurement invariance was assessed for students by class level (first-year vs. senior), parental education (first-generation vs. non-first-generation), sex (male vs. female), and race/ethnicity (White vs. non-White). We selected these groups to examine if the scale adequately performs longitudinally and does not discriminate against historically marginalized populations. To assess measurement invariance, we performed the following steps for each group. First, we tested the fit of the model for each group separately. Next, using generally accepted practices, we tested for configural, metric, and scalar invariance by estimating a series of models and comparing their fit indices incrementally (Putnick & Bornstein, 2016; Rutkowski & Svetina, 2014). After freely estimating parameters for each group, the model for configural invariance confirms an equal number of factors with the same pattern of factor loadings, intercepts, and measurement errors underlie the set of items used

for measuring grit in each group. To test metric invariance, we then constrained the factor loadings to be equal across groups and assessed the magnitude of change in model fit indices to the configural model; very small differences in model fit after constraining parameters provided evidence of metric measurement invariance (see criteria described below). If a model is metric invariant, the construct has the same meaning across the groups examined, allowing for the scales to be used in multivariate statistical analyses. The model for scalar invariance constrained factor loadings to be equal as well as constrained item intercepts across groups. Using model fit change statistics comparing metric to scalar models, we tested whether the groups have equivalent item intercepts. Achieving scalar measurement invariance is necessary for valid group mean comparisons based on latent variables.

We used the following thresholds for assessing measurement invariance as suggested by Chen (2007). For metric (loading) invariance, we used the criteria of a change greater than or equal to $-.010$ in CFI and less than or equal to $.015$ in RMSEA compared to the configural model. The criteria for scalar (loading and intercept) invariance was a change of greater than or equal to $-.010$ in CFI and less than or equal to $.015$ in RMSEA compared to the metric model. We also report the $\Delta\chi^2$; however, due to the large sample size, we relied on the other fit indices to assess measurement invariance.

Concurrent Validity. To assess the concurrent validity of grit, we estimated a series of OLS regression models that predicted NSSE's Engagement Indicators, a perceived gains scale, GPA, and time spent preparing for class using the grit subscales and the control variables described in the data section. Additionally, due to the multi-level structure of our data, we included institution-specific fixed effects in the models. The fixed effects are essentially dummy variables that control for differences in institution attended and encompass both structural

differences in institution type like enrollment size and control and unobserved qualitative factors like a culture of supporting students among faculty or peer support for academics. Additionally, we adjusted the standard errors to account for the clustering of students within institutions. All of the concurrent validity analyses were performed separately for first-year and senior students, following NSSE's standard practices. To aid in the interpretation of the results, we reverse coded the consistency of interest items for the concurrent validity analyses, as the wording of the items in this subscale all indicate lower levels of grit. As we standardized both the outcome variables and grit subscales, the results reported represent the expected standard deviation change in the outcome for a standard deviation change in one of the grit subscales.

Limitations

This study has some limitations that should be taken into consideration when interpreting its results. First, our outcome measures are self-reported by students and may be subject to measurement or social desirability bias. However, self-reported data has generally found to be valid for measures of student behavior and grades, provided that the items on a survey can be comprehended by the respondent, known to the respondent, produces a meaningful response, and the response options are complete and clear (Gonyea 2005). As the Short Grit Scale was administered to students at the same time as the core NSSE instrument, the relationships between grit and our dependent variables should be viewed as correlational, not causal. However, if grit is a personality trait, then we can assume that an individual's level of grit will be relatively constant as the stability of personality traits is one of the most robust findings in psychology, leading towards a greater ability to infer causation (McCrea and Costa Jr. 1994). Additionally, our sample did not have an adequate size to investigate measurement invariance for smaller groups. Therefore, we decided to classify students by race and parental education into two groups

(White/non-White and first-generation/non-first-generation). By aggregating students into larger groups, we may be masking some important variation in our analyses.

Results

We present the results of the validity tests in two sections: construct validity and concurrent validity.

Construct Validity

We began by assessing the model fit of the Short Grit Scale by performing a confirmatory factor analysis. The analysis revealed that the model fit our data marginally, $\chi^2(19)=971.839$ $p < .001$, CFI = .954, TLI = .932, RMSEA = .065 (90% CI .062 - .069). The CFI and TLI values were above our threshold for adequate fit (.90), but the confidence interval for the RMSEA did not include $\leq .06$. Additionally, the standardized factor loading for one item (“Setbacks don’t discourage me”) was extremely low at .10. All other standardized loadings were .59 or greater. The correlation between the two grit subscales was -.51 (the consistency of interest items are reverse coded). The results of the model are visually displayed in Figure 1.

As the fit indices indicated that the fit could be improved and the low loading of the “setbacks don’t discourage me” item (part of the perseverance of effort subscale), we estimated a second CFA without this item. The fit statistics for this analysis were $\chi^2(13)= 630.690$ $p < .001$, CFI = .969, TLI = .951, RMSEA = .064 (90% CI .059 - .068). Removing the low loading item improved the overall model fit, and all of the indices met our thresholds for adequate fit. The standardized loadings for this model were all .59 or greater. The correlation between the subscales was -.52. The standardized item loadings for the revised model are displayed in Figure 2.

Next, we assessed the measurement invariance of the short grit scale using the procedures previously described to examine scale properties across subgroups. Due to the improved fit of the scale without the “setbacks don’t discourage me” item, our analyses were based on the modified model. Table 1 displays the results for baseline models for each group and the combined models testing configural, metric, and scalar invariance.

The baseline model for first-year students indicated adequate fit, while the senior model indicated marginal fit ($RMSEA > .06$). The configural model for class level just met our threshold for adequate fit ($RMSEA$ lower 90% $CI = .060$), indicating the same factor structure form between first-year and senior students (i.e., equivalent number of latent variables, factor loadings, intercepts, and residuals). The change in CFI and RMSEA between the configural and metric models and between the metric and scalar models met Chen’s (2007) suggested thresholds indicating that the modified, two-factor grit model has equivalent loadings and intercepts for both the first-year and senior undergraduate population.

Next, we examined the models by parental education status. The baseline model for first-generation students indicated marginal fit ($RMSEA > .06$) for first-generation students and adequate fit for non-first-generation students. The configural model for parental education met our threshold for adequate fit ($RMSEA$ lower 90% $CI = .059$), indicating that the factor structure form is equivalent between first-generation and non-first-generation students. The change in CFI and RMSEA between the configural and metric models as well as the metric and scalar models indicated relatively little difference in model fit, thus providing evidence for measurement invariance at all levels.

The baseline models for males and females indicated adequate or marginal fit for both groups as both RMSEAs were $> .06$, but the outer edge of the confidence intervals included $.06$.

The configural model also had adequate to marginal fit due to a RMSEA of .064, with a confidence interval of .060 to .068, indicating that the factor structure form is equivalent between male and female students. The change in fit indices between the models also provided evidence for metric and scalar measurement invariance.

Finally, we assessed invariance by race/ethnicity. The baseline model for non-White students had adequate fit. However, the model for White students had marginal fit, RMSEA=.069 (90% CI .064 - .075). The configural model also had a marginal fit due to the RMSEA of .065 (90% CI .061 - .069), indicating equivalent factor structure form. The change indices for the metric and scalar models did not substantially vary. Thus we found measurement invariance at all levels for race/ethnicity.

Concurrent Validity

Due to the findings from the construct validity analyses, we utilized a modified version of the perseverance of effort subscale in our concurrent validity analyses, as the modified subscale was a better fit to the data. Table 2 contains the results of our multivariate analyses. The coefficients represent the expected standard deviation change in the Engagement Indicators, the perceived gains scale, time spent preparing for class, and students grades for a standard deviation change in the two grit subscales consistency of interest and perseverance of effort, holding constant other factors.

Our results indicate that consistency of interest has a negligible or relatively weak association with engagement, perceived gains, time spent preparing for class, and grades after we controlled for other variables. The absolute values of the relationships ranged from .01 for Supportive Environment to .11 for Reflective and Integrative Learning for first-year students. Furthermore, consistency of interest was consistently negatively related to the Engagement

Indicators. However, the relationship was positive for time spent preparing for class and grades. The results were relatively similar for seniors as the absolute value of the coefficients ranged from .02 to .10. However, we did not observe any significant and positive relationships between consistency of interest and our dependent variables for seniors.

In contrast to the consistency of interest results, we found a pattern of positive relationships of a larger magnitude for the perseverance of effort subscale. For first-year students, the perseverance of effort estimates ranged from .15 for Quality of Interactions to .32 for Learning Strategies after holding constant other characteristics. For seniors, the estimates ranged from .11 for Quality of Interactions to .24 for Learning Strategies and Grades. All of the relationships were statistically significant at $p < .001$ for both the first-year and senior subsamples.

Discussion

Admissions leaders are increasingly focusing on using non-cognitive traits, like grit, as a factor in the admissions process (Powell 2013; Sedlacek 2017; Wick 2015). Grit has the theoretical potential to help admissions professionals identify talented students who would thrive in college, yet may lack the traditional academic criteria needed for admission to a highly selective college. Despite the push for using grit as a factor in holistic admissions review, the concept has not been fully empirically validated and is the subject of debate within research circles (Credé et al. 2017; Gonzales-Stokas 2015; Muenks et al. 2017; Ris 2015; Socal 2014). Furthermore, Duckworth's initial validation studies of grit focused on narrow samples atypical of common educational settings (e.g., West Point cadets, Scripps Spelling Bee finalists), indicating the need to reassess the validity of the scale in more representative population (Duckworth et al.

2007; Duckworth and Quinn 2009). Therefore, we sought to investigate the validity of grit and, in particular, the Short Grit Scale, using Messick's (1989) unified framework of validity.

Based on data from nearly 12,000 undergraduates attending a diverse group of U.S. colleges and universities, our study investigated both the construct and concurrent validity of the Short Grit Scale. Though the factor structure of grit has been investigated by others before, we are the first to apply these techniques to a large data set comprised of students attending numerous postsecondary institutions to confirm its factor structure and stability across different student groups. Additionally, grit's relationship to student engagement in educationally beneficial practices has not been extensively studied. Overall, our results demonstrated that 1) a modified, two correlated factor model for grit's consistency of interest and perseverance of effort subscales adequately fits the underlying data; 2) the sub-scale scores generally have the same meaning across different student groups (class level, first-generation status, sex, and racial/ethnic minority status); and 3) the perseverance of effort scale is a more powerful predictor of NSSE measures than consistency of interest, although perseverance of effort's effects are for the most part small in size and vary between class level and across outcomes.

These findings when integrated with prior research allow us to come to some conclusions about grit. First, the Short Grit Scale does not meet generally accepted criteria for use in high-stakes situations. As previously noted by Credé et al. (2017), the proposed factor structure for the Short Grit Scale by Duckworth and Quinn (2009) is invalid (see the Critiques of Grit section for more details). Our confirmatory factor analysis found a poor fit to the data, mainly because one item was relatively unrelated to the perseverance of effort subscale. After removing the item, we achieved adequate fit. However, the scale still did not display good fit, which would be needed to utilize the scale in high stakes decisions like admission to highly selective institutions. However,

it is appropriate for use in research applications, according to this standard. Although, we must also caution that even *if* the scale had properties suited to high-stakes situations, we are concerned that students may not honestly answer the items. Consequently, a preferable approach might be to have counselors rate students on their grit levels like Akos and Kretchmar's (2017) use of the grit scale with family and peers.

Second, we found that one of the grit subscales, perseverance of effort, was significantly and positively related to engagement in educationally purposeful activities, perceived gains, and GPA when we held constant a basket of student characteristics and fixed institutional effects. However, the second subscale was weakly and sometimes negatively related to our dependent variables. Across all outcomes, the average standardized coefficient for perseverance of effort was about .22 and .18 for first-year and senior students, respectively; for consistency of interest, the mean coefficients were .07 and .05, respectively. The grit subscales accounted for a significant proportion of the total explained variance for many of the dependent variables. For instance, for first-year students, the subscales accounted for over 10%, 8%, and 7% of the total variation in learning strategies, perceived gains scales, and GPA, respectively.

Our finding that perseverance of effort, but not consistency of interest, is positively related to a host of outcomes comports with numerous prior studies (e.g., Bowman et al. 2015; Credé et al. 2017; Datu et al. 2015; Muenks et al. 2017). Their relationship to engagement is important due to engagement's association with persistence and student learning (Kuh et al. 2006; National Survey of Student Engagement, n.d.; Pascarella et al. 2010). Furthermore, our measures of engagement were process indicators of student learning, which we would expect students to participate in if grit has positive impacts on college student outcomes. Our concurrent validity results indicate that admissions leaders may want to include students' perseverance of

effort when making admissions decisions in a holistic framework. However, our results do not suggest that perseverance of effort should be a dominant factor in admissions decisions as the partial correlation with our outcomes was not overwhelmingly strong.

Third, our results mainly suggest that grit, when measured by the Short Grit Scale, is invariant across populations. This feature is critical as it suggests that grit is not biased against important subgroups like historically underrepresented racial and ethnic groups, first-generation college students, and women. As grit has been previously critiqued as a vehicle to legitimize class reproduction (Gonzales-Stokas 2015; Ris 2015; Socal 2014), the invariance of grit across these subscales indicates that the critique is not substantiated by the quantitative evidence. Furthermore, it suggests that the use of grit in high stakes decisions, like college admission, should not have a deleterious societal impact, an important consideration in Messick's (1989) validity framework. However, we must be cognizant that grit, as operationalized by the Short Grit Scale, does not meet the standards for use in high stakes decisions; therefore, an invariance analysis should be repeated if an improved scale becomes available.

Given the limited existing research exploring grit's relationship to college student engagement processes, two potential avenues to explore in the future should be considered. First, we should better understand the potential moderating or mediating effects that perseverance of effort has on various student background/demographic characteristics when explaining student engagement behavior and other student outcomes (i.e., academic major, sex, first-generation status, first-generation status). An enticing aspect of grit is its possibility for compensatory effects, where high levels of grit could result in outsized gains for those with lower academic credentials. Second, while Duckworth (2016) has indicated that grit may be malleable, there is limited information about efforts by colleges to increase student grit. Systematic research

exploring who is attempting to change students' levels of grit and how they are going about it can be beneficial for understanding the grit landscape among colleges. Once relevant institutions are identified large-scale data collection efforts can be implemented to track students to see if they are positively impacted by school efforts. Other types of academic success programs can be tracked as well to see how grit-related programs compare to them. Credé and colleagues (2017) suggest that other programs that focus on study skills might be a wiser investment regarding time and funding. Third, our findings in combination with others indicate that the operationalization of grit could be improved (Credé et al. 2017; Muenks et al. 2017). Consequently, future efforts should be made to create an improved grit scale that can conform to strict psychometric scrutiny, given the interest of using non-cognitive factors like grit in high-stakes decisions. Furthermore, much research has demonstrated that the predictive power in grit mainly resides in the perseverance of effort component (Bowman et al. 2015; Credé et al. 2017; Muenks et al. 2017); therefore, such future efforts may want to focus on this component of grit.

Conclusion

Despite the desire of admissions professionals to use grit in their decision-making processes, the current operationalized version of grit does not appear to be a valid measure for high-stakes decisions. However, one dimension of grit, perseverance of effort, shows some promise in its ability to predict important postsecondary outcomes. Additionally, grit appears to be relatively invariant, suggesting that that grit may be a pathway to identify historically underrepresented students that show great promise for postsecondary success. Consequently, more future research should be devoted to creating an empirically valid grit scale.

References

- Akos, P., & Kretchmar, J. (2017). Investigating grit as a non-cognitive predictor of college success. *The Review of Higher Education*, 40(2), 163-186. doi:10.1353/rhe.2017.0000
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association: Washington, DC.
- Bazelaïs, P., David, J.L., & Tenzin, D. (2016). How does grit impact college students' academic achievement in science? *European Journal of Science and Mathematics Education*, 4, 1, 33-43. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1107756.pdf>.
- Bowman, N.A., Hill, P.L., Denison, N., & Bronkema, R. (2015). Keep on truckin' or stay the course? Exploring grit dimensions as differential predictors of educational achievement, satisfaction, and intentions. *Social Psychology and Personality Science*, 6, 639-645. doi: 10.1177/1948550615574300
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling*, 14(3), 464-504.
- Cureton, E. E. (1951). Validity. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 621–694). Washington, D.C.: American Council on Education.
- 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*, 111, 492-511. doi: 10.1037/pspp0000102
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443–507). Washington, D.C.: American Council on Education.

- Cross, T. M. (2013). *Staying the course: Grit, academic success, and non-traditional doctoral students* (Order No. 3587687). Available from ProQuest Dissertations & Theses Global. (1427344572). Retrieved from <https://eric.ed.gov/?id=ED558318>.
- Cross, T. (2014). The gritty: Grit and non-traditional doctoral student success. *Journal of Educators Online*, 11(3), 1-30. Retrieved from <http://www.eric.edu/contentdelivery/servlet/ERICServlet?accno=EJ1033306>.
- Datu, J. F., Valdez, J. P., & King, R. B. (2015). Perseverance counts but consistency does not! Validating the short grit scale in a collectivist setting. *Current Psychology*, 35, 121-130.
- Del Giudice, M. (2014). Grit Trumps Talent and IQ: A Story Every Parent (and Educator) Should Read. Retrieved from <http://news.nationalgeographic.com/news/2014/10/141015-angela-duckworth-success-grit-psychology-self-control-science-nginnovators/>
- Duckworth, A. L. (2013, April). *Angela Lee Duckworth: Grit: The power of passion and perseverance* [Video file]. Retrieved from https://www.ted.com/talks/angela_lee_duckworth_grit_the_power_of_passion_and_perseverance/up-next
- Duckworth, A. L. (2016). *Grit: The power of passion and perseverance*. Simon and Schuster: New York.
- 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(6), 1087-1101. doi: [10.1037/0022-3514.92.6.1087](https://doi.org/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*, 9, 166-174. doi: [10.1080/00223890802634290](https://doi.org/10.1080/00223890802634290)

- Engber, D. (2016). Is “Grit” Really the Key to Success? Retrieved from http://www.slate.com/articles/health_and_science/cover_story/2016/05/angela_duckworth_h_says_grit_is_the_key_to_success_in_work_and_life_is_this.html
- Fosnacht, K. & Sarraf, S. & Howe, E. & Peck, L. K. (2017). How Important are High Response Rates for College Surveys? *The Review of Higher Education*, 40(2), 245-265. [doi:10.1353/rhe.2017.0003](https://doi.org/10.1353/rhe.2017.0003)
- Gonyea, R. M. (2005). Self-reported data in institutional research: Review and recommendations. *New Directions for Institutional Research*, 127, 73-89.
- Gonzalez-Stokas, A. (2015). A genealogy of grit: Education in the new gilded age. *Educational Theory*, 65, 5, 513-528. doi: 10.1111/edth.12130
- Hodge, B., Wright, B., & Bennett, W. (In press). The role of grit in determining engagement and academic outcomes for university students. *Research in Higher Education*.
- Holbein, J. B., Hillygus, D. S., Gibson-Davis, C. M., Hill, D., & Lendard, M. A. (2016). *The development of students' engagement in school, community, and democracy*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2793274.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117-144. doi: 10.1080/03610739208253916
- Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55. doi: 10.1080/10705519909540118
- Kuh, G. D., Kinzie, J., Cruce, T., Shoup, R., & Gonyea, R. M. (2006, July). *Connecting the dots: Multi-faceted analyses of the relationships between student engagement results from the*

- NSSE, and the institutional practices and conditions that foster student success: Final report prepared for Lumina Foundation for Education.* Bloomington, IN: Indiana University, Center for Postsecondary Research.
- MacArthur Foundation. (2017). *Angela Duckworth*. Retrieved from <https://www.macfound.org/fellows/889/>
- McCormick, A. C., Kinzie, J., & Gonyea, R. M. (2013). Student engagement: Bridging research and practice to improve the quality of undergraduate education. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research: Vol. 28.* (pp. 47–92). The Netherlands: Springer.
- McCrae, R. R., & Costa Jr, P. T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, 3, 173-175.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education and Macmillan.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. doi: 10.1037/0003-066X.50.9.741
- Muenks, K. Wigfield, A., Yang, J.S., & O'Neal, C.R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulations, engagement, and achievement. *Journal of Educational Psychology*, 109, 5, 599-620. doi: 10.1037/edu0000153
- National Survey of Student Engagement. (2017). *NSSE's Commitment to Data Quality*. Retrieved from http://nsse.indiana.edu/html/psychometric_portfolio.cfm
- National Survey of Student Engagement. (n.d.). Validity: Predicting retention and degree progress. Retrieved from

http://nsse.indiana.edu/pdf/psychometric_portfolio/PredictiveValidity_retention_degreeprogress.pdf

National Survey of Student Engagement. (2017). *NSSE's Commitment to Data Quality*. Retrieved from http://nsse.indiana.edu/html/psychometric_portfolio.cfm

O'Neal, C. R., Espino, M. M., Goldthrite, A., Morin, M. F., Weston, L., Hernandez, P., & Fuhrman, A. (2016). Grit under duress: Stress, strengths, and academic success among non-citizen and citizen Latina/o first-generation college students. *Hispanic Journal of Behavioral Sciences*, 38, 446-466. doi: 10.1177/0739986316660775

Pace, C. R. (1980). Measuring the quality of student effort. *Current Issues in Higher Education*, 2, 10–16.

Pascarella, E. T., Seifert, T. A., & Blaich, C. (2010). How effective are the NSSE benchmarks in predicting important educational outcomes?. *Change: The Magazine of Higher Learning*, 42(1), 16-22.

Powell, K. (2013). Higher education: On the lookout for true grit. *Nature*, 504, 471–473. doi:10.1038/nj7480-471a

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71-90, doi: 10.1016/j.dr.2016.06.004.

Ris, E. W. (2015). Grit: A short history of a useful concept. *Journal of Educational Controversy*, 10(1), Article 3. Retrieved from <http://cedar.wvu.edu/cgi/viewcontent.cgi?article=1246&context=jec>.

Robinson, W.L. (2015). *Grit and demographic characteristics associated with nursing student course engagement* (Doctoral dissertation, Indiana University Purdue University

- Indianapolis). Retrieved from
https://scholarworks.iupui.edu/bitstream/handle/1805/8008/Robinson_iupui_0104D_10014.pdf?sequence=1
- Rutkowski, L., & Svetina, D. (2017). Measurement invariance in international surveys: Categorical indicators and fit measure performance. *Applied Measurement in Education*, 30(1), 39-51, doi: 10.1080/08957347.2016.1243540.
- Sedlacek, W. (2017). *Measuring Noncognitive Variables: Improving Admissions, Success and Retention for Underrepresented Students*. Sterling, VA: Stylus:.
- Socal, I. (2014). Taking a closer look at the “grit” narratives. *Knowledge Quest*, 43, 1, 8-12.
 Retrieved from <https://eric.ed.gov/?id=EJ1041786>.
- Stewart, S. B. (2015). *Grit and self-control as predictors of first-year student success*. (Doctoral dissertation). Retrieved from
<http://digitalcommons.usm.maine.edu/cgi/viewcontent.cgi?article=1172&context=etd>
- Strayhorn, T. L. (2014). What role does grit play in the academic success of black male collegians at predominately-white institutions?. *Journal of African American Studies*, 18, 1-10. doi: 10.1007/s12111-012-9243-0.
- Suzuki, Y., Tamesue, D., Asahi, K., & Ishikawa, Y. (2015). Grit and Work Engagement: A Cross-Sectional Study. *PLOS ONE*, 10(9), 1-11. doi:10.1371/journal.pone.0137501
- United States Department of Education (2013). *Promoting grit, tenacity, and persistence: Critical Factors for success in the 21st century*. Retrieved from <http://pgbovine.net/OET-Draft-Grit-Report-2-17-13.pdf>.
- Wang, M., & Degol, J. (2016). Staying engaged: Knowledge and research needs in student engagement. *Child Development Perspectives*, 8, 137-143. doi: 10.1111/cdep.12073

West, M. R., Kraft, M.A., Finn, A.S., Martin, R., Duckworth, A.L., Gabrieli, C.F., & Gabriele,

J.D. (2016). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Education Evaluation and Policy Analysis*, 38, 148-170. doi: 10.3102/0162373715597298

Wick, J. (2015). Putting more “grit” in college admissions: Going beyond test optional.

Retrieved from <http://blogem.ruffalonl.com/2015/08/11/putting-grit-college-admissions-test-optional/>.

Wolters, C. A., & Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition and Learning*, 10(3), 293-311. doi: 10.1007/s11409-014-9128-9

Table 1
Fit indices for class level, parental education, sex, and race invariance analyses

Model	Overall fit indices					Comparative fit indices			
	χ^2	df	RMSEA (90% CI)	CFI	TLI	Model comparison	$\Delta\chi^2$	Δdf	ΔCFI $\Delta RMSEA$
<i>Class level</i>									
First-year	237.734***	13	.061 (.054 .068)	0.970	0.951				
Senior	414.245***	13	.066 (.061 .072)	0.967	0.946				
1. Configural	651.979***	26	.064 (.060 .068)	0.968	0.948				
2. Metric	664.530***	31	.059 (.055 .063)	0.967	0.956	1 vs. 2	12.552*	5	0.001 0.005
3. Scalar	753.147***	36	.058 (.055 .062)	0.963	0.957	2 vs. 3	88.617***	5	0.004 0.001
<i>Parental education</i>									
First-generation	352.859***	13	.067 (.061 .073)	0.966	0.946				
Non-first-generation	279.646***	13	.059 (.054 .066)	0.973	0.956				
1. Configural	632.505***	26	.063 (.059 .067)	0.970	0.951				
2. Metric	646.436***	31	.058 (.054 .062)	0.969	0.958	1 vs. 2	13.931*	5	0.001 0.005
3. Scalar	749.836***	36	.058 (.055 .062)	0.964	0.958	2 vs. 3	103.400***	5	0.005 0.000
<i>Sex</i>									
Female	408.30***	13	.064 (.058 .069)	0.969	0.950				
Male	246.177***	13	.065 (.058 .072)	0.969	0.949				
1. Configural	654.479***	26	.064 (.060 .068)	0.969	0.950				
2. Metric	671.091***	31	.059 (.055 .063)	0.968	0.957	1 vs. 2	16.612**	5	0.001 0.005
3. Scalar	829.624***	36	.061 (.058 .065)	0.961	0.954	2 vs. 3	158.533***	5	0.007 -0.002
<i>Race/ethnicity</i>									
Non-white	231.419***	13	.058 (.051 .065)	0.975	0.959				
White	428.660***	13	.069 (.064 .075)	0.964	0.941				
1. Configural	660.079***	26	.065 (.061 .069)	0.968	0.949				
2. Metric	685.467***	31	.060 (.056 .064)	0.967	0.956	1 vs. 2	25.388***	5	0.001 0.005
3. Scalar	793.149***	36	.060 (.057 .064)	0.962	0.956	2 vs. 3	107.682***	5	0.005 0.000

Note: χ^2 = Chi-square; df = degrees of freedom; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; $\Delta\chi^2$ and Δdf = change in χ^2 and degrees of freedom between models; ΔCFI = change in CFI score between models; $\Delta RMSEA$ = change in RMSEA score between models.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 2

Fixed Effect Estimates of the Relationship between Grit and Student Engagement and Perceived Gains for First-Year and Senior Students

	Consistency of interest			Perseverance of effort				
	Est.	Sig.	SE	Est.	Sig.	SE	ΔR^2	Final R^2
<i>First-year students</i>								
Higher-Order Learning	-0.06	**	0.02	0.26	***	0.02	0.06	0.09
Reflective & Integrative Learning	-0.11	***	0.02	0.25	***	0.02	0.05	0.12
Quantitative Reasoning	-0.10	***	0.02	0.22	***	0.02	0.04	0.09
Learning Strategies	-0.02		0.02	0.32	***	0.02	0.10	0.13
Collaborative Learning	-0.09	***	0.02	0.19	***	0.02	0.04	0.09
Discussions w/ Diverse Others	-0.07	***	0.02	0.16	***	0.02	0.02	0.07
Student-Faculty Interaction	-0.05	**	0.02	0.20	***	0.02	0.03	0.07
Quality of Interactions	-0.07	***	0.02	0.15	***	0.02	0.03	0.08
Supportive Environment	0.01		0.02	0.21	***	0.02	0.04	0.06
Perceived gains	-0.06	***	0.02	0.30	***	0.02	0.08	0.11
Time spent: preparing for class	0.08	***	0.02	0.16	***	0.02	0.05	0.09
GPA	0.09	***	0.02	0.26	***	0.02	0.07	0.19
<i>Seniors</i>								
Higher-Order Learning	-0.05	*	0.02	0.17	***	0.02	0.02	0.08
Reflective & Integrative Learning	-0.06	**	0.02	0.17	***	0.02	0.02	0.16
Quantitative Reasoning	-0.07	***	0.02	0.17	***	0.02	0.02	0.13
Learning Strategies	-0.02		0.02	0.24	***	0.02	0.05	0.13
Collaborative Learning	-0.10	***	0.02	0.17	***	0.02	0.03	0.15
Discussions w/ Diverse Others	-0.02		0.02	0.14	***	0.02	0.02	0.07
Student-Faculty Interaction	-0.08	***	0.02	0.19	***	0.02	0.04	0.13
Quality of Interactions	0.00		0.02	0.11	***	0.02	0.02	0.05
Supportive Environment	-0.05	*	0.02	0.16	***	0.02	0.03	0.06
Perceived gains	-0.08	***	0.02	0.20	***	0.02	0.05	0.11
Time spent: preparing for class	0.03		0.02	0.18	***	0.02	0.03	0.10
GPA	0.03		0.02	0.24	***	0.02	0.07	0.23

Note: All dependent variables, consistency of interest, and perseverance of effort were standardized with a mean of 0 and standard deviation of 1. ΔR^2 = change in explained variance after the grit subscales were added to the model. Control variables included race/ethnicity, sex,

standardized test score (SAT/ACT), parental education, major, greek-life participation, age, athletics participation, transfer status, part-time status, educational aspirations, and on-campus residency. Models included fixed institutional effects.

* $p < .05$, ** $p < .01$, *** $p < .001$

¹Modified from Duckworth's proposed subscale. Excludes "setbacks don't discourage me" due to improved model fit without the item (see construct validity section for more details).

Figure 1

Confirmatory Factor Analysis of Duckworth and Quinn's (2009) Short Grit Scale

Note: CI = consistency of interest; PE = Perseverance of effort; See Appendix A for item wordings

Figure 2

Confirmatory Factor Analysis of the Modified Short Grit Scale

Note: CI = consistency of interest; PE = Perseverance of effort; See Appendix A for item wordings

Appendix A

Items in the Short Grit Scale

<i>Variable</i>	<i>Variable Label</i>	<i>Values and labels</i>
GRM1601a	New ideas and projects distract me from previous ones ¹	
GRM1601b	Setbacks don't discourage me.	
GRM1601c	I have been obsessed with a certain idea or project for a short time but later lost interest ¹	1 = Not at all
GRM1601d	I am a hard worker.	2 = Not much
GRM1601e	I set goals but later choose to pursue different ones ¹	3 = Somewhat
GRM1601f	I have difficulty maintaining my focus on projects that take more than a few months to complete ¹	4 = Mostly
GRM1601g	I finish whatever I begin.	5 = Very much
GRM1601h	I am diligent.	

¹Reverse coded in the concurrent validity analyses