



## Grit and conscientiousness: Another jangle fallacy



Annette Ponnock<sup>a,\*</sup>, Katherine Muenks<sup>b</sup>, Monica Morell<sup>a</sup>, Ji Seung Yang<sup>a</sup>, Jessica R. Gladstone<sup>a</sup>, Allan Wigfield<sup>a</sup>

<sup>a</sup> University of Maryland, Human Development & Quantitative Methodology, 3304 Benjamin Bldg., College Park, MD 20742, United States

<sup>b</sup> University of Texas – Austin, College of Education, 1912 Speedway, Stop D5000, Austin, TX 78712, United States

### ARTICLE INFO

#### Article history:

Received 3 April 2020

Revised 20 August 2020

Accepted 10 September 2020

Available online 17 September 2020

#### Keywords:

Motivation/goals

Personality assessment

Adolescent

Advanced quantitative methods

Personality

### ABSTRACT

When grit was first introduced, it gained popularity before basic psychometric questions were fully explored. One critical issue is how distinct grit is from the Big Five personality trait conscientiousness. Most studies have examined correlations between grit and conscientiousness, rather than conducting item-level factor analysis. This study examined the extent to which grit and conscientiousness are empirically distinct, and which predict students' grades. A diverse sample of adolescents completed measures of grit and conscientiousness. MIRT-based confirmatory factor analyses showed that grit and conscientiousness' factor structures strongly overlap. Structural equation modeling showed that conscientiousness and the perseverance of effort component of grit predicted students' grades more strongly than consistency of interest. These findings indicate that grit and conscientiousness are not unique constructs.

© 2020 Elsevier Inc. All rights reserved.

### 1. Introduction

Duckworth and colleagues (Duckworth, Peterson, Matthews, & Kelly, 2007) defined grit as “trait-level perseverance and passion for long-term goals” (p. 1087). They developed measures of grit tapping two components, perseverance of effort (PE) and consistency of interests (CI) and have reported that grit relates to a variety of outcomes. Over the last 10+ years grit has received much attention in research literature, popular press (e.g., Duckworth, 2016), and from educational policymakers. Although there now are many studies of grit (see Credé, Tynan, & Harms, 2017, for a meta-analytic review), fundamental questions regarding its psychometric properties, relations to other similar constructs, and relations to outcomes remain underexplored and/or have produced conflicting results (see Morell et al., in press).

One concern is grit's distinctiveness or overlap with the Big Five trait conscientiousness, the trait from which Duckworth et al. (2007) derived grit. Conscientiousness is defined as being orderly, dependable, and diligent (Costa & McCrae, 1992). Duckworth and colleagues (Duckworth et al., 2007) distinguished the two constructs by arguing that grit concerns individuals' *long-term* rather

than shorter term stamina toward goals and is a “more narrowly defined facet” (Duckworth et al., 2007, p.1088) of personality than conscientiousness. However, when measures of constructs with different names substantially relate/overlap, then a jangle fallacy may be occurring. Marsh et al. (2019) described jangle fallacies as “two scales with apparently dissimilar labels [that] might measure similar constructs” (p. 333). As Credé et al. (2017) put it with respect to grit, “The conceptual similarities between these constructs<sup>1</sup> and grit raises the possibility that the proponents of grit may have fallen victim to... the ‘jangle fallacy’—the belief that two things are different simply because they have different names” (p. 495).

How strongly do grit and conscientiousness relate? In their meta-analysis Credé et al. (2017) reported an overall correlation of  $p = .84^2$  between grit and conscientiousness across 22 studies. Correlations in this range suggest high degrees of association and so may overlap completely. Most research assessing the association between grit and conscientiousness has used correlational analyses at the construct level. Another way to test the strength of association and potential overlap of the two constructs is to conduct factor analyses on both grit and conscientiousness items together to see if the

\* Corresponding author at: The Yale Center for Emotional Intelligence, 350 George St., 3<sup>rd</sup> Floor, New Haven, CT 06511, United States.

E-mail addresses: [annette.ponnock@yale.edu](mailto:annette.ponnock@yale.edu) (A. Ponnock), [kmuenks@utexas.edu](mailto:kmuenks@utexas.edu) (K. Muenks), [mmorell@umd.edu](mailto:mmorell@umd.edu) (M. Morell), [jseyang@umd.edu](mailto:jseyang@umd.edu) (J. Seung Yang), [Jg6314@nyu.edu](mailto:Jg6314@nyu.edu) (J.R. Gladstone), [awigfiel@umd.edu](mailto:awigfiel@umd.edu) (A. Wigfield).

<sup>1</sup> Credé et al. included persistence, proactivity, and conscientiousness in their discussion of conceptually similar constructs.

<sup>2</sup> The  $p$  notation is used to indicate point estimate of correlation obtained from a meta-analysis and to distinguish it from correlations from a single study.

items themselves appear to be indistinguishable to survey participants.

To our knowledge Schmidt et al. (2018) are the only researchers to have factor analyzed items measuring grit with items measuring conscientiousness. They had adolescents and young adults complete a German version of Duckworth and Quinn's (2009) grit scale, and the German version of Costa and McRae's (1992) NEO-PI-R, which measures sub-facets of conscientiousness. Correlations among the facets of conscientiousness and grit ranged from 0.44 to 0.92. They also found that when grit, PE, and CI facets were included in the confirmatory factor analyses, all facets of grit (especially PE) overlapped substantially with the superordinate factor of conscientiousness. Schmidt et al. (2018) concluded that "the results suggest that the perseverance facet of grit cannot be separated from the common factors underlying the NEO PI-R conscientiousness scales" (p. 713). This degree of overlap suggests that the perseverance of effort aspect of grit represents a jangle fallacy.

We advance the correlational work examining the relationship between grit and conscientiousness at the construct level and Schmidt et al.'s (2018) important factor analytic study of overlap/association at the item level in four ways. First, we used multidimensional item response analysis and structural equation models to examine the empirical overlap between grit and conscientiousness. These analyses are broadly housed within the latent variable modeling framework that allows researchers to resolve the problem of attenuated correlations between measurement-error-contaminated variables, among other things (see Morell et al., in press; Rhemtulla, Brosseau-Liard, & Savalei, 2012, for discussion of the advantages of this approach). Second, we test the relative predictive validity of grit and conscientiousness by examining how grit and conscientiousness predicted participants' end of semester grades in math and science classes in high school. Math and science classes become increasingly difficult and therefore students need to be more effortful to succeed in them. Thus, both grit and conscientiousness should relate to this (relatively) long term outcome; to date few studies have examined this prediction. Further, extant work on the relations of grit to achievement often has used the grit total score, even though, as noted above, the PE component of grit relates more strongly to achievement outcomes (Credé et al., 2017). We assessed how the proposed grit subcomponents related to outcomes.

Third, we utilized a large and ethnically diverse U.S. adolescent sample, thus assessing the generalizability of Schmidt et al.'s (2018) findings. Fourth, we examined the factor structure of grit and conscientiousness using a different measure: John et al.'s (1991) conscientiousness scale from their Big Five Inventory (BFI). This frequently used scale in research on personality shows strong convergent validity with Costa and McRae's (1992) NEO-PI-R and is much shorter (nine items versus 64), thus making it ideal for survey research when data collection time is limited and/or respondents are completing a variety of measures (see John & Srivastava, 1999; Soto & John, 2017, for review of work with the BFI and the development of even shorter versions of it).

We examined two main research questions: (1) What is the best-fitting factor structure for grit and conscientiousness? (2) To what extent do the two proposed facets of grit (PE and CI) and conscientiousness predict students' math and science grades? As noted above, we chose math and science grades as the outcome because students tend to view these subjects as difficult. Therefore, we would expect that students' grit and conscientiousness might be particularly strong predictors of their performance in these domains.

We hypothesized that grit and conscientiousness would overlap; that is, items measuring each will load on the same factors, and based on Schmidt et al. (2018), PE items will overlap more with conscientiousness than will CI items. Further, based on work by Muenks, Wigfield, Yang, and O'Neal (2017) and Credé et al.

(2017), we hypothesized that PE, CI, and conscientiousness would predict students' grades, but that PE would relate more strongly to students' grades than would CI.

## 2. Methods

### 2.1. Participants and procedures

Participants were 1,246 9th and 10th grade students (45.7% female) attending seven urban public and independent high schools in the northeastern United States.<sup>3</sup> The mean age of the sample was 14.86 years. The ethnic breakdown was 49% White, 35.7% Black, 6.9% Asian/Asian American, 6.7% Hispanic/Latinx, 0.8% Native American, 0.2% Pacific Islander, and 1.7% another identity (e.g., biracial). We recruited as many participants as possible in order to increase statistical power; no student data was excluded from the analysis. Students took an online or paper survey during school time that contained the measures described below.

### 2.2. Measures<sup>4</sup>

#### 2.2.1. Grit-S

Participants completed the Grit-S developed by Duckworth and Quinn (2009). We chose to use this measure because it has been used most widely in studies of grit. The Grit-S scale contains four items to measure CI and four to measure PE. Participants responded to each item using a 5-point Likert scale (1 = Not at all like me to 5 = Very much like me).

#### 2.2.2. Conscientiousness

Participants completed the 9-item measure of conscientiousness from the Big Five Inventory (BFI) developed by John et al. (1991), which students answered on a 5-point Likert scale (1 = Strongly disagree to 5 = Strongly agree).

#### 2.2.3. Math and science grades

We obtained semester grades in math and science for each participant from school records. Math and science grades were on a 100-point scale ranging from 50 to 100.

Tables S1 and S2 in the Supplemental Materials present the grit and conscientiousness items and reliabilities and bivariate correlations among the latent variables. PE and CI are moderately and significantly correlated ( $r = 0.47$ ), PE and conscientiousness correlated very highly ( $r = 0.87$ ), and CI and conscientiousness are moderately correlated ( $r = 0.53$ ). As predicted PE and Conscientiousness correlated more strongly than did CI and Conscientiousness.

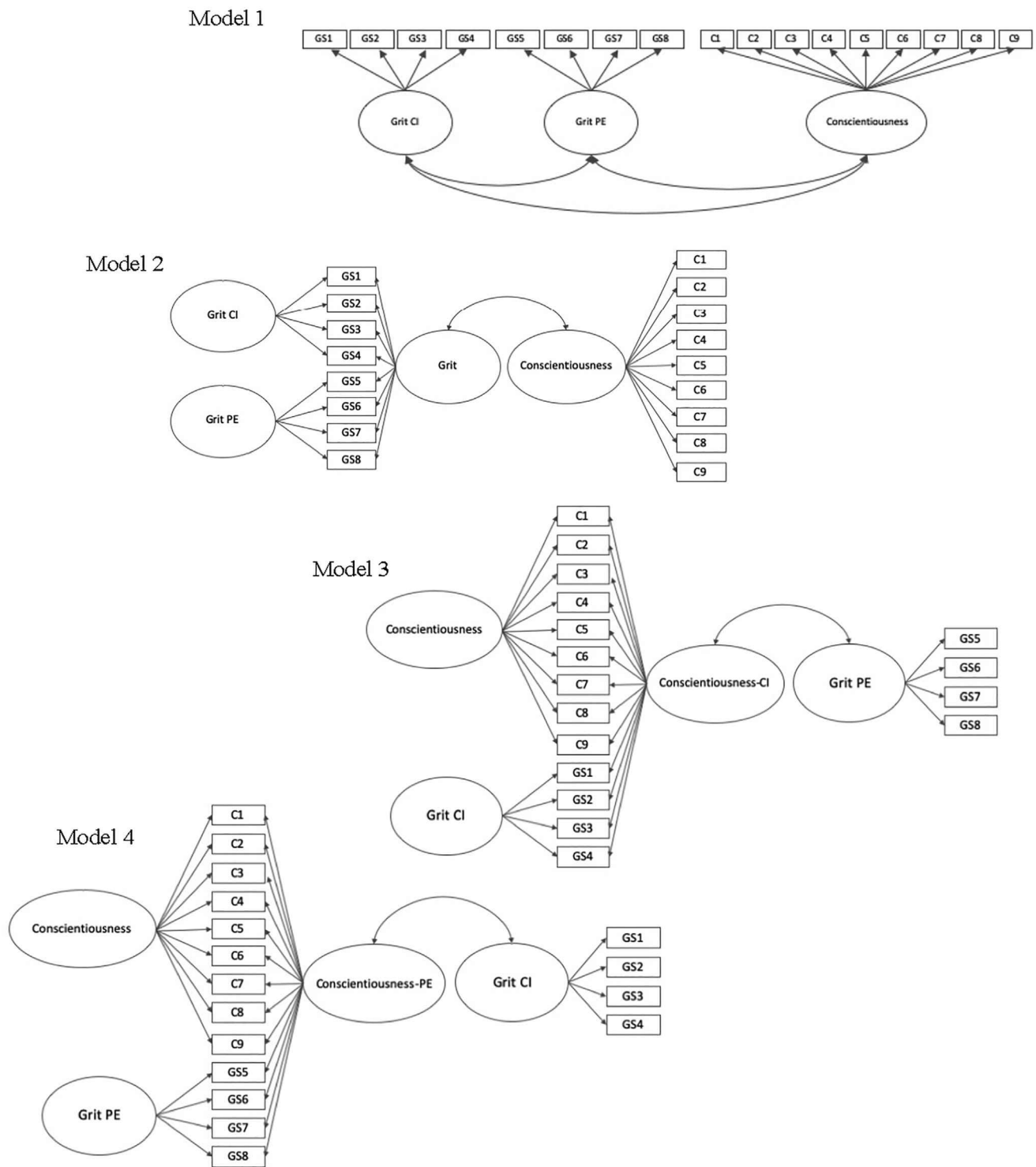
### 2.3. Analysis plan

To answer Research Question 1, we tested four hypothesized models of grit and conscientiousness, using flexMIRT, a computer software package developed for multidimensional item factor analysis with more efficient estimation algorithms (Cai, 2017). The models (see Fig. 1) were: 1) a simple structure model where CI, PE, and conscientiousness items make up separate but correlated factors (Model 1), 2) a bi-factor model (Reise & Revicki, 2014) of grit where CI and PE items make up an overall grit factor that is correlated with a unidimensional conscientiousness factor (Model 2), 3) a bi-factor model of CI and conscientiousness that is correlated with a unidimensional PE factor (Model 3), and 4) a

<sup>3</sup> Students were recruited through their schools via their math, science, or advisory classes. We obtained passive consent from parents and active assent from students. Each student was eligible to receive a \$5 online gift card as a stipend for participating.

<sup>4</sup> These measures were collected as part of a larger study that included measures of motivation, engagement, and identity.

Tested structural models of grit and related constructs.



Note. CI = Consistency of Interest; PE = Perseverance of Effort

Fig. 1. Tested structural models of grit and related constructs.

bi-factor model of PE and conscientiousness that is correlated with a unidimensional CI factor (Model 4).<sup>5</sup> To determine the best fitting

<sup>5</sup> In a bifactor model, items for two constructs (e.g., CI and PE) make up an overall factor (e.g., grit), with the two subfactors (e.g., CI and PE) capturing the residual dependency among the items.

model we used M<sub>2</sub> root mean square error of approximation (RMSEA; Joe & Maydeu-Olivares, 2010), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) fit indices. Models with good fit will have an RMSEA < 0.05 and the lowest AIC, and BIC as compared to other models.

**Table 1**  
Model fit indices and correlations among latent variable from confirmatory item factor analysis of grit and conscientiousness.

	Model 1	Model 2	Model 3	Model 4
-2LL	51268.19	NC	<b>50495.01</b>	50500.87
AIC	51444.19	NC	<b>50693.01</b>	50698.87
BIC	51895.14	NC	<b>51200.33</b>	51206.20
RMSEA	0.03	NC	0.02	0.02
Correlations	$r_{PE \& CI} = 0.47 (0.04)$ $r_{consci \& PE} = 0.87 (0.02)$ $r_{consci \& CI} = 0.53 (0.03)$	-	$r_{consci-CI \& PE} = 0.86 (0.02)$	$r_{consci-PE \& CI} = 0.53 (0.03)$

Note. -2LL =  $-2 * \log$ likelihood; AIC = Akaike information criterion; BIC = Bayesian information criterion; RMSEA = root mean square error of approximation based on  $M_2$ . NC = not converged; PE = Perseverance of effort; CI = Consistency of interest. Consci = Conscientiousness. The best-fitting model indices are bold face.

Finding that Models 1 or 2 fit best would provide evidence that grit and conscientiousness are empirically distinct. Model 1 fitting best also would indicate that the PE and CI components do not make up an overall grit construct, whereas Model 2 fitting best would indicate that two grit components do make up an overall grit construct, as Duckworth et al. (2007) proposed. Support for either Models 3 or 4 will provide evidence for strong empirical overlap of grit and conscientiousness, albeit of different types. Model 3 fitting best suggests that conscientiousness has stronger overlap with CI, whereas Model 4 fitting best suggests that it overlaps more with PE.

For Research Question 2, we ran separate structural equation models for each construct (conscientiousness, PE, and CI) to examine which construct(s) predicted students' math and science grades.<sup>6</sup> We did this for two reasons, one theoretical and one methodological. First, even if grit and conscientiousness overlapped in the factor analyses, they are theoretically distinct and thus we wanted to assess their associations to student outcomes separately in our diverse sample because no one has examined these relations in a sample like ours. Second, given the high correlations between grit and conscientiousness found in other studies (see Credé et al., 2017), we were concerned about multicollinearity, especially between PE and Conscientiousness. Including highly overlapping predictors in the same regression model can result in unreliable standard errors and incorrect statistical inferences when the variance inflation factor (VIF) exceeds 4. VIFs at this level indicate bivariate correlations of 0.85 which we observed in both Models 1 and 3 and so had good reason to be concerned about multicollinearity. We therefore entered each latent variable in the predictive model by itself without controlling for the others. We did however statistically control for gender and ethnicity in all of the predictive models.

### 3. Results<sup>7</sup>

#### 3.1. Research Question 1: What is the best-fitting factor structure for grit and conscientiousness?

Model fit indices are presented in Table 1. The current standard indicating excellent model fit for categorical data is RMSEAs < 0.0125, when using measures with five category responses (Maydeu-Olivares & Joe, 2014). Models 3 and 4 fit the data closest to the excellence criteria and were quite similar to one another, indicating that there is substantial overlap between conscientiousness and CI (Model 3) and conscientiousness and PE (Model 4). All factor loadings were significant. Thus, comparing across the fit indices, Model 3 had marginally better fit than Model

<sup>6</sup> In all of these analyses we treated each item as a categorical rather than continuous variable. Our tests of the distributions of response to the items indicated they were not normally distributed, indicating it is more appropriate to treat the responses as categorical (Rhemtulla et al., 2012).

<sup>7</sup> All data and materials will be shared in the journal's online data repository

4. Model 2 did not converge, suggesting that the two grit factors (CI and PE) do not make up an overall grit construct. Therefore, to test Research Question 2, we examined CI and PE as individual predictors and did not include overall grit in the predictive analyses.

#### 4. Research question 2: To what extent do conscientiousness, PE, and CI predict students' math and science grades?

We next examined how each construct (conscientiousness, CI, and PE) predicted students' math and science grades when entered as single predictors. Conscientiousness, CI, and PE were all significant predictors of both math and science grades at  $\alpha = 0.05$  level. The size of the standardized betas suggests that conscientiousness ( $\beta = 0.22$  and  $0.25$  for math and science, respectively) and PE ( $\beta = 0.25$  and  $0.29$  for math and science, respectively) were the strongest predictors of students' grades, whereas CI ( $\beta = 0.16$  and  $0.15$  for math and science, respectively) was a weaker predictor. Please see Table 3S for standard errors and confidence intervals.

### 5. Discussion

There are two key findings of this study. First, it is the first study conducted in the U.S. that documents the empirical overlap or association of grit and conscientiousness at the item level, thus confirming and extending the work Schmidt et al. (2018) did in Germany with a different conscientiousness measure. This finding also extends previous correlational work on the relations correlation, between grit and conscientiousness at the construct level (see Credé et al., 2017). Based on these results we conclude that grit as measured by the Grit-S is merely a relabeling of the conscientiousness construct that is already established in the literature, and thus a classic jangle fallacy as defined by Marsh et al. (2019).

Second, results of the predictive analyses showed that when examining how PE, CI, and conscientiousness each predict adolescents' STEM achievement, conscientiousness and PE are the strongest predictors. These findings are similar to those of other studies (e.g., Muenks et al., 2017) and buttress Credé et al.'s (2017) conclusion that PE is the main driving force behind the predictive power of grit. One implication of this finding is that researchers using total grit scores in predictive analyses of academic achievement outcomes likely underestimate the strength of those relations due to CI's relatively weak predictive power. Further, given the strong correlation of PE and conscientiousness and their theoretical and empirical overlap, we do not think it is meaningful for researchers to "pit" the two constructs against each other in order to determine which of the two has stronger predictive power.

#### 5.1. Limitations

This study was conducted on high school students from the northeastern United States. Due to these demographics, the find-



ings may not generalize to other populations and/or parts of the world. Acquiescence bias or halo bias could also be an issue with the Grit-S due to the fact that all CI items are negatively worded, and all PE items are positively worded (Billiet & McClendon, 2000). Future research should investigate to what extent this measurement issue could inflate correlations between PE and other constructs measured with positively worded scales.

## 6. Conclusion

Credé (2018), in his essay on grit, asked what we should do about it. We suggest at a minimum that researchers stop measuring it with the Grit-S (see also Morell et al., in press). We propose that either new measures of grit be developed and validated that justify its treatment as a unique construct, or that researchers should stop examining grit and continue to explore conscientiousness in relation to different aspects of achievement. Given the long history of research on conscientiousness and the better measures of it compared to scales frequently used to measure grit we believe the latter approach would be the better of the two.

## Acknowledgements

This research was supported by the National Science Foundation under Grant No. 1534846 to Allan Wigfield and Ji Seung Yang. The authors would also like to thank the school administrators and students who participated in this research, as well as Lara Turci and Amanda Vite for their support with this project.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jrp.2020.104021>.

## References

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling, 7*, 608–628.

- Cai, L. (2017). flexMIRT R version 3.51: Flexible multilevel multidimensional item analysis and test scoring [Computer software]. Chapel Hill, NC: Vector Psychometric Group.
- Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653–665. [https://doi.org/10.1016/0191-8869\(92\)90236-1](https://doi.org/10.1016/0191-8869(92)90236-1).
- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology, 113* (3), 492–511. <https://doi.org/10.1037/pspp0000102>.
- Credé, M. (2018). What shall we do about grit? A critical review of what we know and what we don't know. *Educational Researcher, 47*(9), 606–611.
- Duckworth, A. (2016). *Grit: The power of passion and perseverance*. New York: Scribner.
- 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. <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, 91*(2), 166–174.
- Joe, H., & Maydeu-Olivares, A. (2010). A general family of limited information goodness-of-fit statistics for multinomial data. *Psychometrika, 75*(3), 393–419.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big-Five Inventory-Version 4a and 54*. Berkeley, CA: Berkeley Institute of Personality and Social Research, University of California.
- John, O., & Srivastava, S. (1999). The Big Five taxonomy: History, measurement, and theoretical perspective. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (2nd ed., pp. 102–138). New York: Guilford.
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*, 331–353.
- Maydeu-Olivares, A., & Joe, H. (2014). Assessing approximate fit in categorical data analysis. *Multivariate Behavioral Research, 49*(4), 305–328.
- Morell, M., Yang, J. S., Gladstone, J. R., Faust, L. T., Ponnock, A. R., Lim, H. J., & Wigfield, A. (in press). Grit: The long and short of it. *Journal of Educational Psychology*.
- Muenks, K. M., Wigfield, A., Yang, J. S., & O'Neal, C. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology, 109*, 599–620.
- Reise, S. P., & Revicki, D. A. (Eds.). (2014). *Handbook of item response theory modeling: Applications to typical performance assessment*. Routledge.
- Rhemtulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*, 354–373.
- Schmidt, F. T., Nagy, G., Fleckenstein, J., Möller, J., & Retelsdorf, J. (2018). Same, but different? Relations between facets of conscientiousness and grit. *European Journal of Personality, 32*(6), 705–720. <https://doi.org/10.1002/per.2171>.
- Soto, C. J., & John, O. P. (2017). Short and shorter forms of the Big Five Inventory-2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality, 68*, 69–81.