

Is grit relevant to well-being and strengths? Evidence across the globe for separating perseverance of effort and consistency of interests

David J. Disabato | Fallon R. Goodman | Todd B. Kashdan

George Mason University

Correspondence

Todd B. Kashdan, Department of Psychology, MS 3F5, George Mason University, Fairfax, VA 22030.
Email: tkashdan@gmu.edu

Abstract

Objective: Researchers conceptualize grit as the combination of two facets: perseverance of effort and consistency of interests toward long-term goals. We tested the reliability of grit facet scores across the globe and examined how differently each grit facet related to well-being and personality strengths.

Method: An international sample of 7,617 participants from six of the seven continents (excluding Antarctica) completed an online survey.

Results: Confirmatory factor analyses and omega reliability coefficients indicated that the 12 items from the original Grit Scale were multidimensional and reliably measured perseverance of effort and consistency of interests. Concurrent validity analyses showed that perseverance of effort was moderately to strongly related to subjective well-being, beliefs about well-being, and personality strengths, whereas consistency of interests had weak or negative correlations with these outcomes. The stronger relations with perseverance of effort were replicated across seven regions of the world. The presence of overall grit was supported in individualistic countries, but not collectivistic countries (i.e., those in Latin America and Asia).

Conclusions: We discuss the multidimensionality of grit, including a conceptual understanding of overall grit and how it may differ across cultures. We suggest well-being and strengths researchers study grit facets separately due to their differential validity.

KEYWORDS

grit, perseverance, strengths, well-being

1 | INTRODUCTION

Grit is a personality strength defined by the aggregation of two facets: perseverance of effort and consistency of interests toward long-term goals (Duckworth, Peterson, Matthews, & Kelly, 2007). Perseverance of effort represents sustained effort toward long-term goals despite the presence of setbacks and distress. Consistency of interests represents the passion and dedicated time, attention, and commitment toward long-term goals. Despite intriguing research findings about overall grit (e.g., grittier children outperform peers at the National Spelling Bee; Duckworth, Kirby, Tsukayama, Berstein, &

Ericsson, 2010), the two grit facets have shown unique validity for performance outcomes (Crede, Tynan, & Harms, 2017). The goal of the present study was to assess the reliability of grit facet scores and expand tests of unique validity from performance outcomes to well-being and strengths.

1.1 | Structure and measurement of grit

Grit is typically measured with the Original Grit Scale (Grit-O; Duckworth et al., 2007) or the Short Grit Scale (Grit-S; Duckworth & Quinn, 2009). Both self-report measures contain Likert response scales and two subscales—Perseverance

of Effort and Consistency of Interests—that are typically summed to yield a total grit score. Some measures in psychological science have strongly related facets, thus offering a clear rationale for combining them. For example, the latent correlation between Snyder and colleagues' (1991) Hope Scale facets ranged from .76 to .91 across four samples, and most research with the Hope Scale relies on total scores (Brouwer, Meijer, Weekers, & Baneke, 2008; Gomez et al., 2015). In comparison, the relationship between the two grit facets is significantly weaker. The original exploratory factor analysis on the Grit-O (Duckworth et al., 2007) found a latent correlation of .45 between perseverance of effort and consistency of interests. Subsequent confirmatory factor analyses on the Grit-O and Grit-S have found latent correlations that ranged from .03, .27, .39, to .59 (Abuhassan & Bates, 2015; Datu, Valdez, and King, 2016; Duckworth & Quinn, 2009). The smaller correlations question whether the grit facets should be combined into overall grit scores.

Measurement tools must be able to reliably assess perseverance of effort and consistency of interests—*independent of overall grit*—before researchers can test for the unique validity of the two facets. Haberman (2008) noted that total scores are sometimes more reliable than subscale scores at capturing the facet of interest. For example, Rodriguez, Reise, and Haviland (2016a) found that four of the five subscales from the Multidimensional Anxiety Scale for Children were unreliable and argued they should not be used. If outcome correlation differences are found between two personality facets, but the scores are unreliable, the correlation differences could be due to measurement error. Accordingly, before studying each grit facet separately, the reliability of Grit Scale subscale scores needs to be tested. We are unaware of past research that addressed these important psychometric issues.

Adequate facet score reliability is necessary but insufficient for presenting the advantages of studying grit at the facet level for well-being and strengths. The two grit facets must also show unique validity. Interestingly, the original Grit-O scale development article stated that when predicting performance outcomes, “neither factor was consistently more predictive of outcomes than the other,” suggesting both grit facets were equally important (Duckworth et al., 2007, p. 1091). This conclusion perhaps led researchers to focus on the Grit Scales' total scores rather than subscale scores. Indeed, published research with the Grit Scales usually combines the subscales into a total score. A meta-analysis found 39 academic performance effect sizes using grit total scores, whereas only 11 effect sizes were found using Perseverance of Effort and Consistency of Interests subscale scores (Crede et al., 2017). If perseverance of effort and consistency of interests demonstrate similar concurrent validity for well-being and strengths, then studying each grit facet separately may be less important. However, existing meta-analytic

evidence suggests large effect size differences between the two facets in predicting achievement, retention, and intelligence outcomes (Crede et al., 2017). On average, the perseverance of effort effect sizes were approximately double those of consistency of interests, suggesting the potential utility of studying each facet separately beyond performance outcomes.

1.2 | Well-being, strengths, and grit

Grit has been predominantly studied in the context of achievement and performance outcomes (e.g., Eskreis-Winkler, Shulman, Beal, & Duckworth, 2014), but research suggests grit is also associated with healthy functioning (Hill, Burrow, & Bronk, 2016; Salles, Cohen, & Mueller, 2014; Sheridan, Boman, Mergler, & Furlong, 2015; Singh & Jha, 2008). Although the effect size differences between grit facets when predicting performance outcomes are large, the differences for satisfaction outcomes are even larger (Crede et al., 2017). For example, life satisfaction correlated .41 with perseverance of effort but only .16 with consistency of interests. Correlation patterns with college and job satisfaction were similar. In the present study, we organized healthy functioning outcomes into three groups: subjective well-being (i.e., affect balance and life satisfaction), beliefs about well-being (i.e., beliefs concerning what lifestyles facilitate well-being), and personality strengths (i.e., dispositions that promote psychological adjustment). We hypothesized that perseverance of effort would relate more positively than consistency of interests to every well-being and strengths outcome.

Research on subjective well-being has linked overall grit to life satisfaction, happiness, positive affect, and lower distress (Salles et al., 2014; Singh & Jha, 2008; Vainio & Daukantaitė, 2016). Broader theories of conscientiousness can help explain why perseverance of effort might facilitate subjective well-being. The “invest-and-accrue” model of conscientiousness claims that individuals high on the personality trait invest in behaviors that accrue future success (Hill & Jackson, 2016). Research on older adults suggests that conscientiousness predicts longer life because people invest in their physical health via healthy behaviors (e.g., eating vegetables, walking, flossing; Fry & Debats, 2009). Greater success and goal attainment, when aligned with central life values, then lead to well-being (Sheldon & Elliot, 1999). These mechanisms may be true only for perseverance, as consistency of interests does not correlate as strongly with conscientiousness (Rimfeld, Kovas, Dale, & Plomin, 2016).

Research on beliefs about well-being has found that gritty people are motivated to attain well-being through tasks that promote engagement and meaning in life rather than pleasure (Von Culin, Tsukayama, & Duckworth, 2014). We suspect these effects are driven by perseverance of effort and

not consistency of interests; those who are able to push through adversity may believe pursuing their goals leads to well-being benefits during the journey. The persevering person may believe the unpleasant effort required for long-term goals is worthwhile due to the emergent purpose in life. Maintaining consistent interests over long periods of time does not specify whether the interests are engaging or not. One might even hypothesize that individuals high on consistency of interests are less oriented toward engaging activities because they shy away from trying new things. As for a meaning orientation, perseverance of effort could relate more strongly to altruistic and other-centered goals than consistency of interests.

Research on personality strengths has found that optimism, self-efficacy, gratitude, purpose in life, and goal-directed thinking all exhibit strong correlations with grit (Hill et al., 2016; Sheridan et al., 2015; Vainio & Daukantaitė, 2016). Perseverance of effort may relate more strongly to goal-directed thinking because perseverance directly relates to goal pursuit, whereas consistency of interests does not imply one's interests are tied to a goal. Meaning in life will likely be higher in those who persevere due to the sense of purpose that motivates them. However, consistency of interests could also foster meaning in life through tradition and convention. As for perceptual curiosity, perseverance of effort may be related to curiosity because overcoming setbacks requires a willingness to try new routes and avenues to goal attainment. In contrast, consistency of interests is likely to be inversely related to curiosity given the tendency to stick with the status quo rather than explore new things.

1.3 | Grit across cultures

Most studies on grit have relied on samples from the United States and United Kingdom. The construct of grit has yet to be sufficiently examined beyond the typical use of Western, educated, industrialized, rich, and democratic (WEIRD) samples. This relatively exclusive examination of grit overlooks the possibility that the structure of grit and its relationship with relevant variables may differ in other cultures. Cross-cultural research enables researchers to determine the extent to which personal characteristics or behaviors are basic (i.e., innate to humans) or acquired through experience, in that they arise from basic processes and are embedded in the culture in which a person lives (Haefel, Thiessen, Campbell, Kaschak, & McNeil, 2009). Constructs like grit, which are couched as psychological strengths, may be especially culture-bound because they are embedded in the values and beliefs of a given culture. For example, grit has received criticism for being solely an American ideal, so much so that the *Oxford English Dictionary* (2017) actually classifies the word as “U.S. slang.”

Cross-cultural research on personality finds evidence for the Big Five higher-order personality traits in over 35 countries, including Western, Oceania, Eastern European, Middle Eastern, Asian, and African countries (Allik & McCrae, 2004; Allik & Realo, 2017). Comparatively less research has explored lower-order facets of personality across cultures. While some facets replicate across WEIRD countries, they often fail to replicate in more diverse cultures (e.g., Rollock & Lui, 2016; Thalmayer & Saucier, 2014). Given that grit's facets are at a lower level of personality relative to the broad Big Five domains, we explored cultural differences in both factor structure and the concurrent validity with well-being and strengths.

1.4 | The present study

The goal of the present study was to assess the reliability of grit facet scores and determine the benefits of studying grit at the facet level when examining well-being and strengths. We hypothesized that both perseverance of effort and consistency of interests scores would exhibit adequate reliability, but that overall grit scores would have lower reliability due to a modest association between the two grit facets. We hypothesized that perseverance of effort would have greater concurrent validity than consistency of interests for all three groups of healthy functioning outcomes: subjective well-being, beliefs about well-being, and personality strengths. Given the dearth of research on the cross-cultural differences of grit, we explored for differences in the reliability and validity of grit scores across different regions of the world.

2 | METHOD

2.1 | Participants and procedure

Participants were from the International Wellbeing Study (www.wellbeingstudy.com) between March 2009 and March 2013. The sample consisted of 7,617¹ late adolescent and adult participants (aged 15 years and older²). Age showed a positively skewed distribution, with $M = 33.5$ and $SD = 14.2$. Most participants were women (79%) and married (54%). Participants were compensated with an email summary of score reports, the opportunity to complete online well-being courses, and the chance to win one of fifteen \$100 vouchers. Participants were recruited via snowball sampling from 52 researchers affiliated with the study (the third author being one of them). Researchers were from 22 different countries (representing all continents except Antarctica) and contacted university departments, businesses, charitable organizations, listservs, and online forums for participant recruitment. For example, the third author posted study information on his university department listserv, a positive psychology listserv, and social media accounts, and also asked

journalists to publicize the study in news articles and blog posts. These initial contacts were asked not only to participate, but also to share the study with other people they knew—facilitating the sampling snowball. For example, at the end of each email contact was the following message: “It would help tremendously if you could send this message about our study to your friends/family/colleagues and e-mail lists as well.”

2.2 | Organization into world regions

Participants were from 109 different countries from six of the seven continents (Antarctica excluded). For meaningful comparisons, the 109 countries were grouped into nine world regions. Initial world region groupings were selected according to the *CIA Factbook* (Central Intelligence Agency [CIA], 2015), a reputable source of world information. Past international well-being research has used the *CIA Factbook's* world regions to group international samples (e.g., Tay & Diener, 2011). The *CIA Factbook* organizes countries primarily by geographical proximity, which excludes important contextual information. Categorizations of countries for the current study were made based on additional information about common historical, cultural, and linguistic roots, as well as current economics. For example, due to Mexico's colonial history, Latin culture, Spanish language, and less developed economy, it was separated from North American countries and grouped with South American countries. The United States and Canada were grouped with Northern Europe because of similar ancestral roots and economic development. Southeast Asia was separated from East and South Asia because of greater colonial and Judeo-Christian influence. South Asia was combined with East Asia because of the limited South Asian subsample size ($n = 48$). Participants from Africa and the Middle East were excluded from subsample analyses due to insufficient sample sizes ($n_s = 64$ and 49 , respectively). The final seven world regions were Oceania ($n = 1,964$), Anglo nations and Northern Europe ($n = 1,909$), the former USSR and Eastern Europe ($n = 1,637$), Southern Europe ($n = 874$), Latin America ($n = 719$), East and South Asia ($n = 273$), and Southeast Asia ($n = 138$). Specific countries in each world region are provided in Table S1 in the supporting information.

2.3 | Measures

The International Wellbeing Study assessment battery contained 20 scales (235 items), which took participants an average of 29 min.³ Of these, 12 (sub)scales totaling 88 items were used for the present analyses.⁴ Most participants completed the assessment battery in English (59%); however, the battery was available in 15 other languages. Where no translation was available from the English version to the required

language, scales were translated by a native speaker of that language who had a degree in psychology or higher (most translators were master's or PhD students in psychology familiar with psychometrics). Scales were independently cross-checked after translation by a second translator, and areas of disagreement were identified and resolved between the two translators.

2.3.1 | Grit

The Original Grit Scale (Duckworth et al., 2007) is a 12-item measure of perseverance and passion for long-term goals. The full scale contains two 6-item subscales: Perseverance of Effort (e.g., “Setbacks don't discourage me”) and Consistent Pursuit of Passionate Interests (e.g., “I often set a goal but later choose to pursue a different one” [reverse scored]). Participants responded to items on a 5-point Likert scale ranging from 1 (*not at all like me*) to 5 (*very much like me*). Coefficient alphas for all 12 items ranged from .73 to .82 across the world regions, whereas the perseverance of effort alphas ranged from .71 to .79 and the consistency of interests alphas ranged from .81 to .84.

2.3.2 | Subjective well-being

The Satisfaction With Life Scale (Diener, Emmons, Larsen, & Griffin, 1985) is a five-item measure that assesses one's cognitive appraisal of quality of life (e.g., “I am satisfied with my current life”). Items are rated on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Coefficient alphas ranged from .89 to .91 across world regions.

The Subjective Happiness Scale (Lyubomirsky & Lepper, 1999) is a four-item measure of global happiness. The first item is a rating of absolute happiness, and the second item assesses happiness relative to one's peers. The third and fourth items ask participants how well characteristics of happiness describe them (e.g., “Some people are generally very happy, and they enjoy life regardless of what is going on, getting the most out of everything”). All items are rated on 7-point Likert scales. The first two item responses range from 1 (*less happy*) to 7 (*more happy*), and the last two item responses are from 1 (*not at all*) to 7 (*a great deal*). Coefficient alphas ranged from .74 to .86 across the world regions.

The Center for the Epidemiological Studies Depression Scale (Radloff, 1977) is a 20-item measure of depression containing four factors: depressed affect, positive affect, somatic symptoms, and interpersonal distress (Shafer, 2006). An example item is “I felt that everything I did was an effort.” Items are rated on a 4-point Likert scale ranging from 0 (*rarely or none of the time*) to 3 (*most or all of the time*). Coefficient alphas ranged from .89 to .92 across the world regions.

2.3.3 | Beliefs about well-being

The Orientations to Happiness Scale (Peterson, Park, & Seligman, 2005) is an 18-item measure that consists of three 6-item subscales representing three different personal beliefs about how to seek well-being: a Pleasure Orientation (e.g., “Life is too short to postpone the pleasures it can provide”), developed from hedonic theories of well-being; an Engagement Orientation (e.g., “I am always very absorbed in what I do”), developed from psychological flow; and a Meaning Orientation (e.g., “I have a responsibility to make the world a better place”), developed from eudaimonic theories of well-being. Participants responded to items on a 5-point Likert scale ranging from 1 (*not at all like me*) to 5 (*very much like me*). Across the world regions, coefficient alphas ranged from .75 to .81 for the Pleasure Orientation subscale, .54 to .73 for the Engagement Orientation subscale, and .72 to .84 for the Meaning Orientation subscale.

2.3.4 | Personality strengths

The Adult Hope Scale (AHS; Snyder et al., 1991) is an eight-item measure of goal-directed thinking containing two subscales: Agency, or goal-directed will (e.g., “I meet the goals I set for myself”), and Pathways, or goal-directed flexibility (e.g., “I can think of many ways to get the things in life that are important to me”). Only the Pathways subscale was included due to semantic overlap between the Agency subscale and the Perseverance of Effort subscale of the Grit Scale. Items are rated on an 8-point Likert scale ranging from 1 (*definitely false*) to 8 (*definitely true*). Coefficient alphas ranged from .70 to .81 across the world regions.

The Curiosity and Exploration Inventory-II (Kashdan et al., 2009) is a 10-item measure of the motivational system associated with the pursuit of novelty and challenge. The measure consists of two 5-item subscales: Stretching, or seeking out new knowledge and experiences (e.g., “I actively seek as much information as I can in new situations”), and Embracing, or willingness to embrace novelty and uncertainty (e.g., “I am the type of person who really enjoys the uncertainty of everyday life”). Items are rated on a 5-point Likert scale ranging from 1 (*very slightly or not at all*) to 5 (*extremely*). Coefficient alphas ranged from .84 to .90 across the world regions.

The Meaning in Life Questionnaire (Steger, Frazier, Oishi, & Kaler, 2006) is a 10-item measure of the presence of and search for meaning in life. Only the five-item Presence subscale was used, which assesses the extent to which individuals view their lives as meaningful (e.g., “I understand my life’s meaning”). Items are rated on a 7-point Likert scale ranging from 1 (*absolutely untrue*) to 7 (*absolutely true*). Coefficient alphas ranged from .85 to .92 across the world regions.

3 | RESULTS

3.1 | Analytic strategy

We conducted a series of three confirmatory factor analyses (CFA) on the total sample: (a) one-factor model of general grit, (b) two-factor model of grit’s facets, and (c) bifactor model (see Figure 1). Conceptually, the two-factor model assumes items reflect only one construct and that each construct could be studied in isolation. A bifactor model assumes items reflect multiple constructs, both a broad trait and a specific facet, which are hierarchically organized. The bifactor CFA model is conceptually similar to a hierarchical CFA, but it is advantageous for modeling multidimensional constructs because it allows for significance tests of the general factor or specific factors above and beyond the other(s) (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012). We allowed for respecification of the measurement models with theoretically justifiable cross-loadings or correlated errors. If a two-factor model adequately fit the data, factor rho reliability coefficients would be calculated (Raykov, 1997). If a bifactor model adequately fit the data, omega reliability coefficients would be calculated to ensure the subscale scores⁵ were reliable (Rodriguez, Reise, & Haviland, 2016b). Following reliability analyses, we tested the concurrent validity of perseverance of effort and consistency of interests scores on the total sample. Latent correlations between the grit factors and subjective well-being, beliefs about well-being, and personality strength were estimated. To determine the degree of misinterpretation that could result from using Grit-O total scores, we also calculated observed correlations.

All analyses were replicated across seven world regions. We used a multiple-group CFA to test for measurement invariance of the best-fitting total sample CFA model across world regions (Bryne, 2016). Given our interest in only the covariance structure of the data (i.e., not the mean structure), we only tested for weak measurement invariance (i.e., factor loadings constrained to be equal). If full weak invariance was not achieved, partial weak measurement invariance was tested for. The same reliability coefficients and latent and observed correlations were calculated within each world region.

3.2 | Confirmatory factor analyses

Total sample descriptive statistics (i.e., means, standard deviations, skewness, and kurtosis) describing the 12 grit items are presented in Table 1. To test the reliability of perseverance of effort and consistency of interests scores, confirmatory factor analyses were conducted via structural equation modeling in the lavaan R package (Rosseel, 2012). Because the grit items were measured on an ordinal scale, and sufficiently normally distributed (i.e., no item skew coefficients

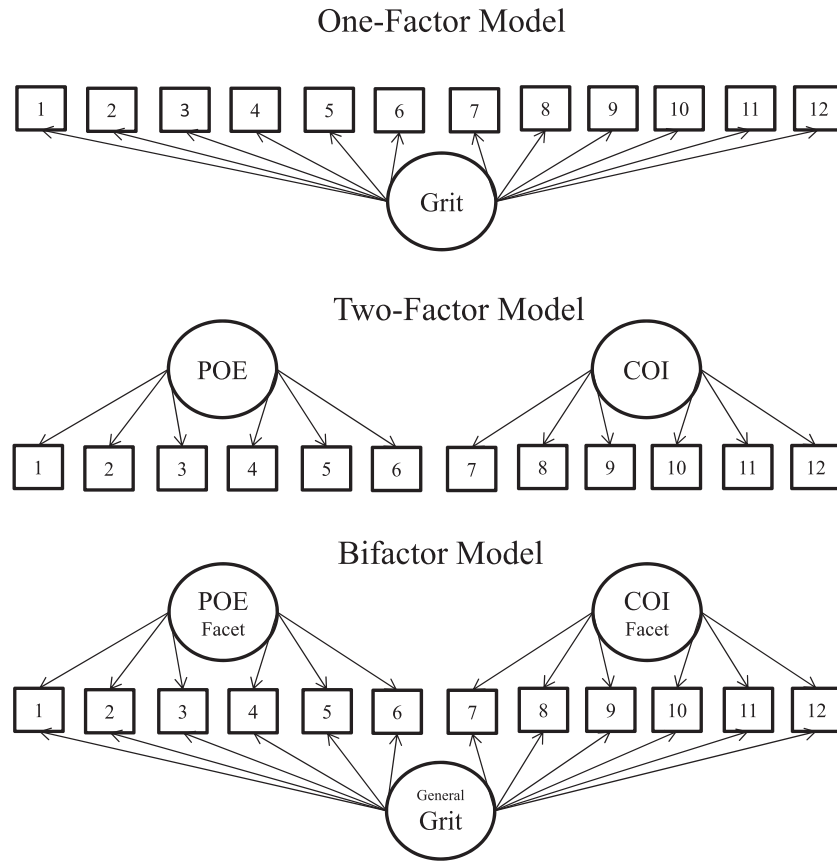


FIGURE 1 The three confirmatory factor analysis models tested. POE = perseverance of effort; COI = consistency of interests

TABLE 1 Item descriptive statistics

Subscale	Item	Wording	Mean	SD	Skew	Kurtosis
Perseverance of Effort	1	I have overcome setbacks to conquer an important challenge.	3.62	1.07	-0.49	-0.45
	2	Setbacks don't discourage me.	3.03	1.12	-0.06	-0.81
	3	I am a hard worker.	3.90	1.04	-0.79	-0.01
	4	I finish whatever I begin.	3.44	1.14	-0.42	-0.72
	5	I have achieved a goal that took years of work.	3.63	1.32	-0.63	-0.80
	6	I am diligent.	3.73	1.06	-0.58	-0.35
Consistency of Interests	7	New ideas and new projects sometimes distract me from previous ones.	2.86	1.14	0.09	-0.77
	8	My interests change from year to year.	3.47	1.16	-0.45	-0.64
	9	I have been obsessed with a certain idea or project for a short time but later lost interest.	3.42	1.23	-0.41	-0.80
	10	I often set a goal but later choose to pursue a different one.	3.55	1.14	-0.47	-0.60
	11	I have difficulty maintaining my focus on projects that take more than few months to complete.	3.72	1.24	-0.67	-0.60
	12	I become interested in new pursuits every few months.	3.45	1.19	-0.37	-0.80

Note. SD = standard deviation. Kurtosis coefficients were centered around zero. Consistency of interest item descriptives are shown after reverse coding.

TABLE 2 Measurement model fit indices

Model	χ^2	df	Scaled χ^2	Scaled df	CFI	NNFI	RMSEA	WRMR
Null	94,293.95	66	30,253.82	21.17				
One-factor	17,888.32	54	12,145.53	23.07	.599	.632	.263	12.53
Two-factor	6,231.10	53	5,535.25	29.72	.818	.870	.156	7.39
Two-factor with three correlated errors	4,581.76	50	4,330.55	28.76	.858	.895	.140	6.34
Bifactor	1,920.09	42	2,838.36	30.04	.907	.935	.111	4.10
Bifactor with one correlated error	1,125.20	41	1,615.93	28.48	.947	.961	.086	3.14
Configural (i.e., unconstrained) invariance	1,459.70	287	1,727.73	168.80	.947	.956	.093	3.58
Partial weak invariance for specific factors	1,813.46	359	1,152.55	146.69	.966	.968	.080	3.99
Partial weak invariance for general factor	2,293.21	359	903.29	113.23	.973	.967	.081	4.49
Full weak invariance for all factors	2,789.58	431	976.25	143.10	.972	.973	.074	4.95

Note. χ^2 = chi-square value; CFI = confirmatory fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual factor. All chi-square values were significant at $p < .001$.

or kurtosis coefficients exceeded ± 1.0), polychoric correlations rather than Pearson product-moment correlations were used⁶ (Holgado-Tello, Chacón-Moscoso, Barbero-García, & Vila-Abad, 2010). Any zero values in the 5×5 bivariate contingency tables used to calculate the polychoric correlations were constrained to .001 for estimation (Savalei, 2011). Although not standardized per se, factor loadings estimated with polychoric correlations can be interpreted as the latent correlation between the item and factor. A robust version⁷ of the diagonally weighted least squares (DWLS) estimator is appropriate for fitting CFAs with ordinal items and was therefore used (Flora & Curran, 2004).

Conventional model fit indices were used to compare the model fit of the CFA models. The comparative fit index (CFI), non-normed fit index (NNFI; also known as the Tucker-Lewis index), root mean square error of approximation (RMSEA), and weighted root mean square residual (WRMR) were used to assess model fit (Hu & Bentler, 1999). The CFI and NNFI range from 0 to 1.0, with .95 or greater indicating adequate model fit. RMSEA ranges from 0 to 1.0, with .08 or less indicating adequate model fit. WRMR ranges from 0 to ∞ , with lower values indicating better model fit (DiStefano, Liu, Jiang, & Shi, 2018). Although there is less simulation research on the WRMR, preliminary evaluation suggests a value of 1.0 or less indicates adequate model fit. For the scaled chi-square test statistic, lower values indicate better model fit. A scaled chi-square difference test ($\Delta\chi^2$) was used to compare models (Satorra & Bentler, 2001). Because this test is more likely to be significant with large samples, the difference in each model fit index was also used to compare model fit, when the chi-square scaling factors were similar across models.

3.3 | One- and two-factor models

The one-factor model with no correlated errors resulted in poor fit (see Table 2). The fit was too poor for attempts at respecification and was thus rejected. The two-factor model with no cross-loadings or correlated errors also resulted in poor fit. A scaled chi-square difference test showed that the two-factor model fit the observed data better than the one-factor model, $\Delta\chi^2(1) = 463.45$, $p < .001$. Evaluation of item content and modification indices were used to determine whether model fit could be improved. An alternative model with three correlated errors was tested. The first correlated error was Items 3 and 6 due to their similar item content about work ethic ($r = .29$). The second was Items 8 and 12 due to their similar item content about having new interests ($r = .25$). The third was Items 1 and 2 due to their similar item content about setbacks ($r = .19$). Even with the three correlated errors, the model fit was still unacceptable.

3.4 | Bifactor model

The bifactor model with no correlated errors resulted in fair, but still unacceptable fit. A scaled chi-square difference test showed that the bifactor model fit the observed data better than the two-factor model with three correlation errors, $\Delta\chi^2(8) = 869.53$, $p < .001$. The CFI increased .049, the NNFI increased .040, the RMSEA decreased .029, and the WRMR decreased 2.24.⁸ Evaluation of item content and modification indices were used to determine whether model fit could be improved. One correlated error between Items 3 and 6 was added ($r = .32$), which significantly decreased the chi-square value, $\Delta\chi^2(1) = 857.64$, $p < .001$. The correlated

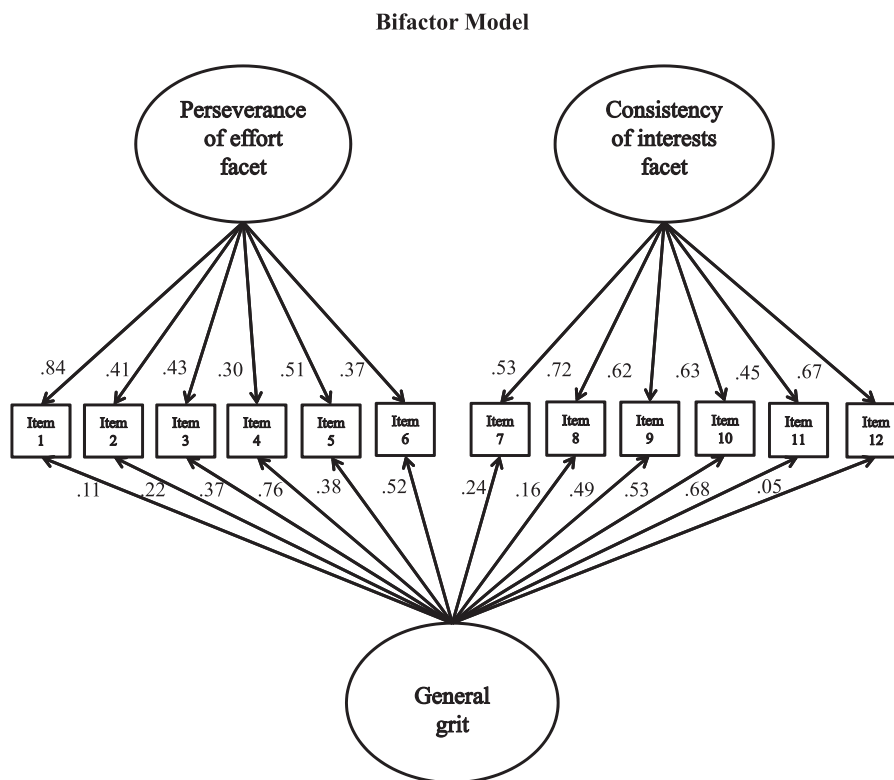


FIGURE 2 Factor loadings from final bifactor model with the total sample

error likely reflects the grit nuance (i.e., personality attribute below facets) of work ethic under the facet of perseverance of effort (Mottus, Kandler, Bleidorn, Reimann, & McCrae, 2017). Although adding in correlated errors between items 8 and 12 as well as Items 1 and 2 from the modified two-factor model significantly decreased the chi-square value, $\Delta\chi^2(2) = 36.69, p < .001$, the model fit indices remained or worsened: The CFI stayed the same, the NNFI decreased .001, the RMSEA increased .001, and the WRMR increased 0.61. Therefore, it is unclear whether new interests and setbacks nuances are present or whether the residual item correlations are simply due to sampling error. For the sake of parsimony, we kept out the two additional correlated errors. With the one-error correlation, the model fit reached acceptable levels, with the CFI and NNFI close to or above .95 and the RMSEA close to .08. However, the WRMR was not close to the 1.0 desired value. Figure 2 presents the factor loadings from this final model.

3.5 | Reliability

Although the bifactor model best fit the data, this result does not necessarily imply the Grit Scale is truly multidimensional (Rodriguez et al., 2016b). Any use of observed subscale scores from a bifactor model should be preceded by tests of reliability. Omega coefficients⁹ are recommended for assessing the reliability of total and subscale scores from

multidimensional measures and can be directly calculated from the results of a bifactor model. As a requirement for bifactor models, the latent correlations between the general grit factor, perseverance of effort specific factor, and consistency of interest specific factor were constrained to zero, which allows the sources of item variance to be divided into three groups: variance due to the general factor, variance due to the specific factors, and variance unique to the item. The proportion of total score variance explained by the general factor is the omega hierarchical coefficient. The proportions of subscale score variance explained by each specific factor are the omega hierarchical subscale coefficients. The uniqueness is the proportion of total or subscale score variance not due to any factor and is interpreted as error variance.¹⁰

The acceptable reliability cut-offs for alpha coefficients cannot necessarily be applied to omega hierarchical and omega hierarchical subscale coefficients. For example, an omega hierarchical subscale coefficient less than .70 does not imply that grit subscale scores are unreliable. The issue of reliability is more complicated because there are multiple sources of reliable variance. A suggested cut-off for determining the reliability of observed scores can come from the explained common variance. The explained common variance is the ratio of variance explained by a target factor over the total variance explained by any factor.¹¹ If the explained common variance is above .60, then we interpret observed scores to sufficiently capture the target factor (e.g., Brouwer,

TABLE 3 Total sample outcome correlations

Outcome Type of correlation	General grit		Perseverance of effort		Consistency of interests		Facet correlation difference	
	Latent	Observed	Latent	Observed	Latent	Observed	Latent	Observed
Subjective happiness	.26	.31	.32	.37	.02	.14	.30*	.23*
Life satisfaction	.26	.24	.20	.29	-.02	.10	.22*	.19*
Depression	-.27	-.34	-.19	-.29	-.13	-.24	-.06	-.05
<i>Average correlation</i>	.26	.30	.24	.32	.04	.16	.19	.16
Engagement orientation	.30	.27	.43	.45	-.18	.00	.61*	.45*
Meaning orientation	.19	.24	.48	.42	-.13	-.01	.61*	.43*
Pleasure orientation	.00	-.10	.20	.13	-.34	-.26	.54*	.39*
<i>Average correlation</i>	.16	.14	.37	.33	-.22	-.09	.59	.42
Goal-directed flexibility	.26	.29	.51	.48	-.15	.00	.66*	.48*
Perceptual curiosity	.20	.13	.46	.41	-.34	-.17	.80*	.58*
Meaning in life	.31	.41	.40	.45	.09	.21	.31*	.24*
<i>Average correlation</i>	.26	.28	.46	.45	-.13	.01	.59	.43

Note. The sign of the depression correlations was reversed for calculating the average correlations. Any correlation .04 or greater in magnitude is significant at $p < .001$. Significance test ($*p < .001$) for facet correlation difference was Steiger's z-test (1980).

Meijer, & Zevalkink, 2013; DeSousa et al., 2014). Therefore, explained common variances above .60 for the Perseverance of Effort and Consistency of Interests subscale scores are needed to psychometrically justify the use of Grit Scale subscales. As with any statistical cut-off, .60 should be interpreted as a rough guideline rather than a strict rule (Lance, Butts, & Michels, 2006).

Using the final bifactor model (with one correlated error), the omega hierarchical and omega hierarchical subscale coefficients were calculated (see the first row of Table 4). For all 12 grit items, the explained common variance of the general grit factor was $.43/ (.43 + .45) = .49$, which suggests less than half the reliable total score variance is due to the general factor. For the six Perseverance of Effort items, the explained common variance of the specific perseverance of effort factor was $.48/ (.48 + .32) = .60$, which suggests most of the reliable subscale score variance is due to the perseverance of effort factor. For the six Consistency of Interests items, the explained common variance of the consistency of interests factor is $.64/ (.64 + .23) = .74$, which suggests most of the reliable subscale score variance is due to the consistency of interests factor. The greater variance explained by the specific factors compared with the general factor support the use of Grit Scale subscale scores. In sum, the reliability statistics suggest the Grit Scale is a truly multidimensional measure with no one source of variance dominating (e.g., general factor).

3.6 | Concurrent Validity Analyses

After confirming the reliability of grit facet scores, we moved to testing the concurrent validity of each to subjective well-being, beliefs about well-being, and personality strengths (Cronbach & Meehl, 1955). We computed latent correlations between the grit facet factors in the final bifactor model (with one correlated error) and nine observed outcome scores: three subjective well-being measures, three beliefs about well-being measures, and three personality strengths measures (see Table 3). Hypothesis tests (i.e., Steiger's z-test) were conducted to determine significant differences between the perseverance of effort and consistency of interests correlations for each outcome variable (Meng, Rosenthal, & Rubin, 1992). We also computed correlations between the nine outcomes and the general grit factor. Because the most common practice with the Grit Scale is to use average (or summed) total scores, we also calculated observed correlations with the Grit Scale scores and the nine observed outcome scores. The average correlation¹² rows indicate the average correlation across each group of outcomes.

All perseverance and consistency of interests latent correlations were significantly different from one another, with a correlation difference of at least .22, except for depression. The perseverance of effort latent correlations with subjective well-being outcomes were moderate in size, whereas the consistency of interests latent correlations were near zero, except for depression. The perseverance of effort latent correlations

TABLE 4 Omega reliability coefficients across world regions

World region Statistic	Sample statistics		General grit		Perseverance of effort		Consistency of interests	
	Subsample size	Observed facet correlation	Omega hierarchical	Explained common variance	Omega hierarchical subscale	Explained common variance	Omega hierarchical subscale	Explained common variance
Total sample	7,617	.23***	.43	.49	.48	.60	.64	.74
Oceania	1,964	.20***	.42	.48	.51	.63	.63	.72
Anglo nations and Northern Europe	1,909	.24***	.44	.50	.46	.56	.65	.75
Former USSR and Eastern Europe	1,637	.33***	.46	.52	.45	.54	.64	.73
Southern Europe	874	.23***	.39	.46	.50	.63	.66	.76
Latin America	719	.11**	.37	.43	.47	.59	.71	.83
East and South Asia	273	.02	.32	.38	.46	.56	.75	.91
Southeastern Asia	138	.19*	.43	.49	.47	.57	.66	.75

Note. The perseverance of effort and consistency of interests columns are based upon half of the Grit Scale items. * $p < .05$. ** $p < .01$. *** $p < .001$.

with beliefs about well-being were large for an engagement and meaning orientation and more moderate for a pleasure orientation. The consistency of interests latent correlations were all significantly negative, especially for a pleasure orientation. The perseverance of effort latent correlations with personality strengths were all large, with the largest association occurring for goal-directed flexibility. The consistency of interests latent correlations were all less than .10, with significant negative correlations for goal-directed flexibility and perceptual curiosity. The observed correlations made perseverance of effort and consistency of interests both appear more related to subjective well-being. In general, the observed correlations for consistency of interests were positively biased, likely because the observed scores included some of the general grit factor.

The correlations with the general grit factor suggest overall grit is associated with subjective well-being and personality strengths. The only outcome overall grit was not related to was a pleasure orientation to well-being. Several of the observed correlations for general grit were positively biased, likely because the total scale scores contained some of the perseverance of effort factor variance. All the correlations together suggest the potential for misinterpreting the influence of consistency of interests. Only general grit and perseverance of effort are associated with greater subjective well-being, beliefs about well-being, and personality strengths.

3.7 | World region analyses

A series of multiple-group CFAs of the final bifactor model was conducted to determine the degree of measurement

invariance across world regions (see Table 2). In each model, the correlated error was allowed to differ across world regions. The model fit of the configural invariance (i.e., all factor loadings allowed to vary across world regions) model was marginal, likely due to its lack of parsimony. The full weak invariance model (i.e., all factor loadings fixed across world regions) fit significantly worse¹³ than the configural invariance model, $\Delta\chi^2(144) = 193.62$, $p = .004$. We proceeded to test for partial weak invariance for the specific and general factors separately. The partial weak invariance model for the specific factors did not fit significantly worse than the configural invariance model, $\Delta\chi^2(72) = 79.01$, $p = .267$. The partial weak invariance model for the general factor fit marginally worse than the configural invariance model, $\Delta\chi^2(72) = 89.72$, $p = .077$. The two partial weak invariance models had the same degrees of freedom and thus no chi-square difference test could be computed. Because the partial weak invariance model for the specific factors fit the data better than that for the general factor (and the general factor loadings were marginally variant), the partial weak invariance model for the specific factors was used for reliability and concurrent validity analyses in each world region.

Using the partial weak invariance model for the specific factors (see Figures S1-S7), the item variance explained by the general grit factor, variance explained by the specific facet factors, and variance unique to the items were calculated. These results were used to calculate omega hierarchical, omega hierarchical subscale, and the explained common variances (see Table 4). In general, results are consistent with the total sample results. The reliability of the grit facet

scores is greater than the reliability of the overall grit scores for every world region, except for the former USSR and Eastern Europe. This pattern was particularly evident in the Latin America and East and South Asia world regions, which may reflect the weak observed facet correlations (see column 2 of Table 4). With regard to the grit facets, the consistency of interests scores were more reliable than the perseverance of effort scores. While the consistency of interests explained common variances were all above .70, most of the perseverance of effort ones were below the .60 cut-off.

To test the differential concurrent validity of each grit facet across world regions, perseverance of effort and consistency of interests latent and observed correlations were compared (see Tables S2–S8). The average correlations for subjective well-being, beliefs about well-being, and personality strengths are reported in Table S9. For subjective well-being, the average perseverance of effort latent correlations were approximately double that or more of consistency of interests in all world regions, except Southeastern Asia. For beliefs about well-being and personality strengths, the average perseverance of effort latent correlations were moderate to strong and positive, whereas the average consistency of interests correlations were weak and negative, leading to large facet correlation differences in each world region. In all world regions, the observed Grit Scale total score correlations were more similar to the perseverance of effort correlations than the consistency of interests correlations.

4 | DISCUSSION

Most research on grit has focused on overall grit, but recent research suggests studying each grit facet separately offers unique validity for performance outcomes. We sought to extend this claim to well-being and strengths outcomes by examining the concurrent validity of the two grit facets. First, we tested whether each grit facet could be reliably measured by the Grit Scale, an issue ignored in past research. Omega coefficients and explained common variance based off a final bifactor model suggested perseverance of effort and consistency of interests scores were more reliable than overall grit scores. Latent and observed correlations with subjective well-being, beliefs about well-being, and personality strength outcomes showed unique concurrent validity for each grit facet. For every outcome (excluding depression), perseverance of effort had stronger positive correlations than consistency of interests. In general, the observed overall grit correlations were more similar to the perseverance of effort correlations, potentially hiding the weak or negative relationships between consistency of interests and healthy functioning. Accordingly, well-being and strengths researchers are encouraged to study grit at the facet level.

4.1 | The multidimensionality of grit

The factor analysis and reliability results confirm the true multidimensionality of grit. Although some “multidimensional” constructs in psychology are actually unidimensional and belong to a single general factor (Rodriguez et al., 2016a), grit does not appear to be one of them. A one-factor model fit the data very poorly, and the observed correlation between grit facet scores was modest ($r = .23$). Reise, Morizot, and Hays (2007) suggest that when facets are correlated less than .40, the specific factors tend to be meaningfully distinct from any general factor. Furthermore, a two-factor model did not fit the data well, even after considering several correlated errors. The Grit-O does a better job of assessing facets independent of their broader trait than common personality measures (Rodriguez et al., 2016a). Across 50 different multidimensional measures, the average omega hierarchical subscale reliability coefficient of personality facet scores was .27 compared with .48 and .64 for perseverance of effort and consistency of interest scores in the present study. Therefore, one benefit of the Grit-O is that researchers can be confident subscale score effects are not solely due to general grit and effect differences are not solely due to measurement error.

Truly multidimensional constructs are difficult to measure with observed scores because the reliable variance is split between multiple sources of variance, which in this case is general grit and the two grit facets. With this in mind, it is still hard to justify the use of Grit Scale total scores when the explained common variance of .49 indicates the grit facet factors explained more item variance than the general grit factor. The average omega hierarchical reliability coefficient was .80 across 50 different multidimensional measures (Rodriguez et al., 2016a), a value significantly larger than the .43 found for the Grit-O in the present study. With a correlation between grit facets greater than .40, the reliability of Grit-O total scores should be larger. With the hopes of obtaining more reliable scores for overall grit, researchers should consider using the Short Grit Scale (i.e., Grit-S). In the final bifactor model, six of the seven items that had general factor loadings greater than .35 are included in the eight-item Grit-S (Duckworth & Quinn, 2009). Consistent with this suggestion, a meta-analysis found the latent correlation between the two Grit-S subscales ($r = .66$) to be much larger than that of the Grit-O subscales ($r = .27$; Crede et al., 2017). Importantly, the Grit-O and Grit-S are not the only way to measure overall grit, and alternative measures might do a better job. Of course, latent variable modeling (e.g., bifactor model in the present study) could also be used to capture overall grit.

4.2 | Overall grit

Some researchers have claimed that general grit does not “exist”¹⁴ due to the unique concurrent validity of the grit

facets (Crede et al., 2017). We disagree with this reasoning because distinct correlations are expected for lower-order facets embedded within hierarchical personality structures, such as the Five-Factor Model of personality. If the facets lacked sufficient discriminant validity, there would be no reason to propose them. For example, facets of Agreeableness show unique concurrent validity for aggression, with correlations ranging from $-.13$ to $-.50$ (Crowe, Lynam, & Miller, 2017), and yet most psychological scientists believe Agreeableness “exists.” Rather than distinct facet correlations, factor analysis is needed to determine the dimensionality of a construct (Chen et al., 2012). Our factor analyses provide evidence for a general grit factor in most world regions, even though perseverance and consistency of interests showed divergent observed and latent correlations.

Of course, dimensionality alone is insufficient to justify the scientific study of a construct. A philosophical issue at the heart of factor analysis is whether the extracted factors make conceptual sense and/or have practical utility. A general factor can be extracted from any two constructs that share variance—even an observed correlation as low as $.10$. Especially with a weaker association between them, combining two constructs must have a relatively clear meaning and prove fruitful for scientific inquiry. We believe that overall grit passed this conceptual evaluation. Grit is the ability to continue pursuing the same long-term goal despite setbacks or other opportunities. Some researchers have reduced grit down to simply work ethic (Crede et al., 2017). However, Duckworth et al. (2007) uses a basic example to illustrate the difference: “Consider two children learning to play the piano . . . [T]hese children are matched in the intensity of effort they expend toward musical training . . . [the one] who practices intensively yet moves from piano to the saxophone to voice will likely be surpassed by an equally gifted but grittier child [who sticks with piano].” Diligence alone is not enough to distinguish between these two children. Persevering toward *the same long-term goal* is grit.

Grit might be best studied in the context of a specific activity, interest, or goal. For example, Robertson-Kraft and Duckworth (2014) tested whether overall grit predicted teacher retention in public schools. Does a teacher in a low-income school district have the grit to stay at his or her school and not transfer to a high-income school district (i.e., consistency of interests), even if he or she receives pay cuts and is not appreciated by students’ parents (i.e., perseverance of effort)? Researchers could modify the Grit Scale to ask individuals about a particular activity, interest, or goal similar to the goal-specific Hope Scale (Feldman, Rand, & Kahle-Wroblewski, 2009). Grit seems worthy of study when trying to understand why people persist with a particular activity, interest, or goal in their lives (e.g., Olympic athletes).

One can still imagine a dispositional tendency for effortful pursuits of the same long-term goals across life domains (e.g.,

career, marriage, spirituality, recreation). But how many different long-term goals can a person have? Studies suggest that people have multiple life goals, but many of them are shorter term (Emmons, 1992), and it is unclear whether people have multiple purposes across life domains or a single overarching life aim (McKnight & Kashdan, 2009). If people only have one or two goals at a time that span several years, trait grit would make little conceptual sense. Qualitative research into the number of long-term goals endorsed by children and adults may clarify the merit of overall grit as a global disposition (and the adjective *gritty*).

4.3 | Perseverance of effort

Perhaps researchers who have used the Grit Scales to measure overall grit as a trait disposition are actually interested in measuring perseverance of effort. For example, overall grit was theorized to be a buffer against Internet addiction (Maddi et al., 2013). Conceptually, perseverance of effort would be relevant for overcoming addiction given the inevitable relapses common; however, consistency of interests not so much. If the participant completing the Grit Scale was responding to items specifically for the long-term goal of overcoming addiction, then consistency of interests would be relevant. But as a trait disposition, consistency of interests refers to any and all interests, projects, and goals. Ironically, consistency of interests might be higher for someone who is struggling with Internet addiction because he or she would maintain interest in Internet use! Maddi et al. (2013) found that hardiness was a better predictor of Internet addiction than grit, but readers are left wondering whether perseverance of effort alone would have been equivalent to hardiness.

The positive association between perseverance of effort with well-being and strengths likely reflects the broader construct of Conscientiousness. Angela Duckworth, the creator of the Grit Scale, has suggested perseverance is a facet of Conscientiousness (MacCann, Duckworth, & Roberts, 2009). Although perseverance of effort is not in Costa and McCrae’s (1995) six facets of Conscientiousness, the facets of the Five-Factor model are far from solidified. Roberts and colleagues (Roberts, Bogg, Walton, Chernyshenko, & Stark, 2004; Roberts, Chernyshenko, Stark, & Goldberg, 2005) have identified different facets (e.g., industriousness, self-control, and orderliness) depending on what measures are included. However, not every facet of Conscientiousness contributes to well-being. In one study, the facet of dutifulness only correlated $.03$ with subjective well-being (Quevedo & Abella, 2011). Our study highlights the importance of lower-order facets for understanding personality (Paunonen & Ashton, 2001).

Future work should examine mechanisms of perseverance of effort that are independent of overall conscientiousness. Hope theory offers a likely mechanism for perseverance

(Snyder, 2002). Hope theory proposes successful goal attainment leads to subjective well-being, particularly positive emotions, whereas unsuccessful goal attainment leads to distress and negative emotions. An important caveat is that attaining long-term goals that are externally motivated and inconsistent with one's values may not have well-being benefits (Sheldon & Elliot, 1999). For long-term goals, successful goal attainment would include incremental victories along the way, such as one's child receiving an A on a third-grade quiz or one's child completing seventh grade, on the way to the long-term goal of helping one's child graduate from high school. In this way, perseverance of effort is similar to agency in Snyder's (2002) hope theory, defined as the belief that one is capable of working toward goal attainment despite the presence of obstacles and setbacks. In fact, the Agency subscale of the Adult Hope Scale is approaching tautology with the Perseverance of Effort subscale of the Grit Scale (and thus removed from the present study). For example, one Perseverance of Effort item is "I finish whatever I begin" and one Agency item is "I meet the goals that I set for myself." Researchers interested only in perseverance may consider using Snyder's Hope Scales instead of the Grit Scale.

4.4 | Consistency of interests

Our findings raise questions about why consistency of interests is unrelated—or even inversely related—to aspects of well-being and personality strengths. Simply put, feeling passionate about the same interests year after year may not be a route to well-being and strengths. Possessing any long-term goal that evokes passion (even if this goal changes) may be more important than keeping the same long-term goal. A college student might change her passion and interests from academics during the school year to physical fitness in the summer, while still persevering at each goal. Her well-being is unlikely hampered by switching passions or interests; instead, this switch might signal a healthy form of psychological flexibility (Kashdan & Rottenberg, 2010). Hope theory is agnostic about whether a person changes his or her goals over time, which may reflect why goal-directed thinking predicts well-being more strongly than consistency of interests (Magaletta & Oliver, 1999).

Alternatively, the negligible association of consistency of interests with well-being and strengths may reflect deficiencies in the Grit Scale measure more so than the construct of consistent passion over time. Passion is defined as "a strong inclination toward an activity that people like, that they find important, and in which they invest time and energy" (Vallerand et al., 2003, p. 757). Upon examination of individual items, the Consistency of Interests subscale might be measuring a lack of novelty-seeking or even rigidity more so than passion (e.g., "I become interested in new pursuits every few months" [reverse scored]). This explanation is partially

supported by the inverse correlations that consistency of interests exhibited with perceptual curiosity and a pleasure orientation to happiness. In comparison, items from Vallerand and colleagues' (2003) Harmonious Passion subscale,¹⁵ which consistently relate to well-being, include "I am completely taken with this activity" and "The new things that I discover with this activity allow me to appreciate it even more." Future researchers might consider generating new consistency of interests items to more closely align with conceptualizations of passion.

While the Consistency of Interests items in the Grit-O may not measure the sustained passion for long-term goals that Duckworth et al. (2007) conceptually introduced, the subscale scores clearly measure a personality trait. The consistency of interests factor had the largest loadings of any factor in the final bifactor model, and consistency of interests had a moderate, negative association with depression, suggesting it may be relevant to anhedonia symptoms. Perhaps consistency of interests, as measured by the Grit Scales, should be thought of as a unique construct completely distinct from grit and its facets of perseverance and passion.

4.5 | Cross-cultural differences

In general, findings in each region of the world were similar to the total sample results, yet each analysis showed a few discrepant world regions. In regard to the structure of grit, most regions of the world with individualistic cultures—Oceania, Anglo nations and Northern Europe, the former USSR and Eastern Europe, and Southern Europe—showed similar Grit Scale score reliability. The reliability of grit total scores (omega hierarchical) for Latin America¹⁶ and East and South Asia were lower than other world regions, resulting from the .11 and .02 observed correlations between the Perseverance of Effort and Consistency of Interests subscale scores. Such little overlap between grit facets in these world regions suggests the Grit Scale does a poorer job of assessing overall grit in collectivistic countries. Perhaps individuals from collectivistic cultures are motivated to persevere as a way to contribute to their families and communities, regardless of their level of personal interest or investment. For example, in a study of Japanese participants, grit was positively correlated with an orientation toward altruistic meaning, but not individual engagement (Suzuki, Tamesue, Asahi, & Ishikawa, 2015). Surprisingly, the total score reliability for Southeastern Asia was equivalent to the total sample value, with an observed subscale correlation of .19. A sample of participants from the Philippines, a country included in the present study's Southeastern Asia world region, generated a latent correlation of only .03 between the two grit facets (Datu et al., 2016). These results are in direct contrast to one another and may be due to sampling error from the smaller subsample size in the present study. Based on our findings,

researchers should be particularly cautious about using Grit-O or Grit-S total scores in Latin American or Asian regions of the world and other countries with collectivistic cultures.

General grit factor loadings were more variable across world regions than the perseverance and consistency factor loadings, suggesting cross-cultural differences in what exactly overall grit is (see Figures S1–S7). The largest differences were found for the world regions with collectivistic cultures: Latin America, East and South Asia, and Southeastern Asia. Grit may be characterized by consistent striving for a personal goal that can be accomplished (e.g., becoming president of a company) in individualistic countries, whereas grit may be characterized by personal goals that are more relational or growth oriented (e.g., be a better parent) in collectivistic countries. Although the perseverance of effort explained common variances were similar across world regions, those for consistency of interests were considerably larger for the Latin America and East and South Asia world regions. Essentially for these two world regions, consistency of interests items did not load as strongly on the general grit factor; most of the item variance was explained by the consistency of interests factor. It begs the question whether overall grit was even measured or whether simply two different aspects of perseverance were captured in the perseverance of effort items (e.g., work ethic and overcoming obstacles). The consistency of interests items may fail to incorporate collectivist forms of interest and passion for long-term goals, potentially indicating a form of cultural bias.

With regard to concurrent validity (see Tables S2–S9), results for the Oceania, Anglo nations and Northern Europe, and Southern Europe world regions were very similar. Unlike with the structure of grit, discrepancies were found for the former USSR and Eastern Europe world region. Overall grit related less strongly to subjective well-being, beliefs about well-being, and personality strength outcomes. Grit also had smaller facet correlation differences, particularly for subjective well-being, likely due to the greater overlap between perseverance of effort and consistency of interests (observed subscale correlation of .33). Participants from the former USSR and Eastern European countries may have interests more tied to their work than people in other countries, strengthening the connection between the two grit facts. The Latin America, East and South Asia, and Southeastern Asia facet correlation differences were similar to the total sample results, suggesting the greater relevance of perseverance of effort across the globe. The main exception was depression, where consistency of interests was equally or more relevant. Depression may be characterized more so by anhedonia than sadness in these parts of the world. Of course, with all world region discrepancies, effects could be due to cross-cultural limitations of the Grit Scale as opposed to real construct differences. Future research with a variety

of measures assessing grit and various facets of Conscientiousness can help tease these two possibilities apart.

5 | LIMITATIONS

Several limitations warrant mentioning. First, this study relies upon retrospective, self-report measures, which are prone to several biases (Stone et al., 2002). Future research on grit would benefit from multiple-informant reports, which provide more valid measures of a construct than single reports (Vazire, 2006). Second, it is unclear how representative the sample is within each country and world region. Although participants were from 109 different countries, the sampling method was snowball rather than randomization. One consequence was a large proportion of women compared with men in the sample. Results cannot necessarily be interpreted as normative. Third, the translation of the self-report measures used in the present study is not immune to systematic measurement error that could have biased findings. Without in-depth cognitive interviews, it is impossible to know whether item semantics were consistent across translations. While multiple qualified translators were used for each language, back-translation could have enhanced accuracy (Beaton, Bombardier, Guillemin, & Ferraz, 2000).

6 | CONCLUSION

Grit has become one of the most popular constructs in today's media, highlighting the need for researchers to understand its components. Our findings suggest grit researchers interested in studying well-being and strengths should report separate results for perseverance of effort and consistency of interests due to acceptable subscale score reliability and substantial prediction differences between the two facets. Depending on the research question and culture of participants, overall grit may or may not be meaningful.

ACKNOWLEDGMENT

The author(s) received no financial support for the research, authorship, and/or publication of this article.

CONFLICT OF INTERESTS

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

ENDNOTES

¹ The final sample size reflects participants who completed all measures. Eighty-two percent of participants who started the online survey completed all measures. Unfortunately, demographic information comparing

completers and noncompleters was not available because demographic questions were asked at the end of the survey.

- ² Our total sample contained people (8.9%, $n = 679$) who were less than 18 years old. The majority of adolescents were from Finland, Slovenia, the Philippines, Greece, and New Zealand. Principal Investigator Aaron Jarden received ethical permission to include late adolescents from the Open Polytechnic of New Zealand Ethics Committee, September 2009.
- ³ Participants who completed the online survey in less than 15 min were excluded due to assumed careless responding, resulting in 7,617 final participants.
- ⁴ Copies of all measures and items used in the International Wellbeing Study are available at the study website: <http://www.wellbeingstudy.com>. Click on the “Findings” tab and scroll down to the Study Question section. Measures and items from each language used are available.
- ⁵ Reliability and validity are characteristics of scores for specific populations and purposes (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014). A measure, such as the Grit Scale, is not universally reliable or valid. Rather, our analyses seek to test the reliability and validity of the Grit Scale scores with online survey completers for assessing general grit, perseverance of effort and consistency of interests.
- ⁶ By forcing participants to choose discrete response options with Likert scale items, Pearson product-moment correlations are attenuated. Polychoric correlations disattenuate the correlations, thus violating fewer statistical assumptions, creating greater statistical power, and generating better model fit than Pearson product-moment correlations.
- ⁷ We used a mean- and variance-adjusted robust estimator that does not change the parameter estimates, but rather scales the model chi-square value, degrees of freedom, and standard errors due to the inefficiency of the diagonally weighted least squares estimator (Savalei, 2014).
- ⁸ Note that the scaling factors across all total sample models were similar (i.e., ranging from 0.65 to 0.75), allowing us to compare model fit indices.
- ⁹ Coefficient alpha is typically used to assess the reliability of measures, but it is inappropriate for multidimensional measures because it assumes reliable variance comes from one source (Cortina, 1993). With multidimensional measures, the reliable variance is due to multiple sources.
- ¹⁰ Technically, the uniqueness of an item includes two sources of variance: that specific to the item and random measurement error. Although not necessarily error variance, specific variance is often merged with error variance because there is no way to separate out the two sources of item variability.
- ¹¹ Traditionally, the explained common variance is only applied to the general factor (Rodriguez et al., 2016a); however, because we were particularly interested in the specific factors, we applied the explained common variance principle to the specific factors.
- ¹² Because depression is inversely related to subjective well-being, the opposite sign of the coefficients was used to calculate the averages.
- ¹³ The scaling factors greatly differed across the measurement invariance models (range = 0.845 to 2.857). Therefore, some of the scaled chi-square values were smaller for worse-fitting models. The original chi-square values were used to obtain the parameter estimates and are the

only statistically valid comparison across models because the scaled chi-square values are not on the same metric across models (Savalei, 2014). Note that the model fit indices are still based off the scaled chi-square values and scaled degrees of freedom and thus cannot be compared across models with different scaling factors. However, model fit indices can be used to accept or reject a model in isolation.

- ¹⁴ *Exist* is considered in an empirical sense of the word, not necessarily ontological, as we can never truly know if psychological constructs exist in reality (Slaney & Garcia, 2015).
- ¹⁵ Vallerand and colleagues' (2003) Passion scale contains two subscales: Harmonious and Obsessive Passion. Harmonious passion is the autonomous drive to engage in activities a person likes; obsessive passion is the uncontrollable urge to engage in activities a person likes. While harmonious passion relates to greater well-being, obsessive passion is associated with less; therefore, we emphasize the Harmonious Passion subscale items.
- ¹⁶ The omega hierarchical reliability coefficients for Latin America and Southeastern Asia are slight overestimates, as Item 12 had a negative general grit factor loading in both world regions (see Figures S5 and S7).

ORCID

Todd B. Kashdan  <http://orcid.org/0000-0001-6438-0485>

REFERENCES

- Abuhassan, A., & Bates, T. C. (2015). Grit: Distinguishing effortful persistence from conscientiousness. *Journal of Individual Differences, 36*, 205–214.
- Allik, J., & McCrae R. R. (2004). Toward a geography of personality traits: Patterns of profiles across 36 cultures. *Journal of Cross-Cultural Psychology, 35*, 13–28.
- Allik, J., & Realo, A. (2017). Universal and specific in the five factor model of personality. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 173–190). Oxford, England: Oxford University Press.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: AERA Publications.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *SPINE, 25*, 3186–3191.
- Brouwer, D., Meijer, R. R., Weekers, A. M., & Baneke, J. J. (2008). On the dimensionality of the Dispositional Hope Scale. *Psychological Assessment, 20*, 310–315.
- Brouwer, D., Meijer, R. R., & Zevalkink, J. (2013). On the factor structure of the Beck Depression Inventory–II: G is the key. *Psychological Assessment, 25*, 136–145.
- Bryne, B. M. (2016). Adaption of assessment scales in cross-national research: Issues, guidelines, and caveats. *International Perspectives in Psychology: Research, Practice, Consultation, 5*, 51–65.
- Central Intelligence Agency (2015). The world factbook. Retrieved from <https://www.cia.gov/library/publications/the-world-factbook/>.
- Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J. P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted

- constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, 80, 219–251.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98–104.
- Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 64, 21–50.
- Crede, 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, 492–511.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Crowe, M. L., Lynam, D. R., & Miller, J. D. (2017). Uncovering the structure of agreeableness from self-report measures. *Journal of Personality*. <https://doi.org/10.1111/jopy.12358>.
- Datu, J. A. D., Valdez, J. P. M., & King, R. B. (2016). Perseverance counts but consistency does not! Validating the Short Grit Scale in a collectivist setting. *Current Psychology*, 35, 121–130.
- DeSousa, D. A., Zibetti, M. R., Trentini, C. M., Koller, S. H., Manfro, G. G., & Salum, G. A. (2014). Screen for child anxiety related emotional disorders: Are subscale scores reliable? A bifactor model analysis. *Journal of Anxiety Disorders*, 28, 966–970.
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71–75.
- DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, 25, 453–466.
- Duckworth, A. L., Kirby, T., Tsukayama, E., Berstein, H., & Ericsson, K. (2010). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science*, 2, 174–181.
- 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*, 9, 1087–1101.
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT-S). *Journal of Personality Assessment*, 91, 166–174.
- Emmons, R. A. (1992). Abstract versus concrete goals: Personal striving level, physical illness, and psychological well-being. *Journal of Personality and Social Psychology*, 62, 292–300.
- Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, 5, 1–12.
- Feldman, D. B., Rand, K. L., & Kahle-Wroblewski, K. (2009). Hope and goal attainment: Testing a basic prediction of hope theory. *Journal of Social and Clinical Psychology*, 28, 479–497.
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9, 466–491.
- Fry, P. S., & Debats, D. L. (2009). Perfectionism and the five-factor personality traits as predictors of mortality in older adults. *Journal of Health Psychology*, 14, 513–524.
- Gomez, R., McLaren, S., Sharp, M., Smith, C., Hearn, K., & Turner, L. (2015). Evaluation of the bifactor structure of the Dispositional Hope Scale. *Journal of Personality Assessment*, 97, 191–199.
- Haberman, S. J. (2008). When can subscores have value? *Journal of Educational and Behavioral Statistics*, 33, 204–229.
- Haefffel, G. J., Thiessen, E. D., Campbell, M. W., Kaschak, M. P., & McNeil, N. M. (2009). Theory, not cultural context, will advance American psychology. *American Psychologist*, 64, 570–571.
- Hill, P., Burrow, A., & Bronk, K. (2016). Persevering with positivity and purpose: An examination of purpose commitment and positive affect as predictors of grit. *Journal of Happiness Studies*, 17, 257–269.
- Hill, P. L., & Jackson, J. J. (2016). The invest-and-accrue model of conscientiousness. *Review of General Psychology*, 20, 141–154.
- Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, 44, 153–166.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The Curiosity and Exploration Inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, 43, 987–998.
- Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, 30, 865–878.
- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, 9, 202–220.
- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, 46, 137–155.
- MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences*, 19, 451–458.
- Maddi, S. R., Erwin, L. M., Carmody, C. L., Villarreal, B. J., White, M., & Gundersen, K. K. (2013). Relationship of hardiness, grit, and emotional intelligence to Internet addiction, excessive consumer spending, and gambling. *Journal of Positive Psychology*, 8, 128–134.
- Magaletta, P. R., & Oliver, J. M. (1999). The hope construct, will, and ways: Their relations with self-efficacy, optimism, and general well-being. *Journal of Clinical Psychology*, 55, 539–551.
- McKnight, P. E., & Kashdan, T. B. (2009). Purpose in life as a system that creates and sustains health and well-being: An integrative, testable theory. *Review of General Psychology*, 13, 242–251.
- Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111, 172–175.
- Mottus, R., Kandler, C., Bleidorn, W., Riemann, R., & McCrae, R. R. (2017). Personality traits below facets: The consensual validity,

- longitudinal stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology*, *112*, 474–490.
- Oxford English Dictionary. (2017). Grit [Def. 2]. In *Oxford English Dictionary Online*. Retrieved from <https://en.oxforddictionaries.com/definition/grit>
- Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*, 524–539.
- Peterson, C., Park, N., & Seligman, M. E. (2005). Orientations to happiness and life satisfaction: The full life versus the empty life. *Journal of Happiness Studies*, *6*, 25–41.
- Quevedo, R. J. M., & Abella, M. C. (2011). Well-being and personality: Facet-level analyses. *Personality and Individual Differences*, *50*, 206–211.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*, 329–353.
- Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*, 19–31.
- Rimfeld, K., Kovas, Y., Dale, P., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of Personality and Social Psychology*, *111*, 780–789.
- Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality*, *38*, 164–178.
- Roberts, B. W., Chernyshenko, O. S., Stark, S. E., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on several major personality questionnaires. *Personnel Psychology*, *58*, 103–139.
- Robertson-Kraft, C., & Duckworth, A. L. (2014). True grit: Trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. *Teachers College Record*, *116*, 1–27.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*, 223–237.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*, 137–150.
- Rollock, D., & Lui, P. P. (2016). Measurement invariance and the five-factor model of personality: Asian international and Euro American cultural groups. *Assessment*, *23*, 571–587.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36.
- Salles, A., Cohen, G. L., & Mueller, C. M. (2014). The relationship between grit and resident well-being. *American Journal of Surgery*, *207*, 251–254.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.
- Savalei, V. (2011). What to do about zero frequency cells when estimating polychoric correlations. *Structural Equation Modeling*, *18*, 253–273.
- Savalei, V. (2014). Understanding robust corrections in structural equation modeling. *Structural Equation Modeling*, *21*, 149–160.
- Shafer, A. B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology*, *62*, 123–146.
- Sheridan, Z., Boman, P., Mergler, A., & Furlong, M. J. (2015). Examining well-being, anxiety, and self-deception in university students. *Cogent Psychology*, *2*, 1–17.
- Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*, 482–497.
- Singh, K., & Jha, S. D. (2008). Positive and negative affect and grit as predictors of happiness and life satisfaction. *Journal of the Indian Academy of Applied Psychology*, *34*, 40–45.
- Slaney, K. L., & Garcia, D. A. (2015). Constructing psychological objects: The rhetoric of constructs. *Journal of Theoretical and Philosophical Psychology*, *35*, 244–259.
- Snyder, R. C. (2002). Hope theory: Rainbows of the mind. *Psychological Inquiry*, *13*, 249–275.
- Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., ... Harney, P. (1991). The wills and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, *60*, 570–585.
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, *53*, 80–93.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251.
- Stone, A. A., Turkman, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (2002). *The science of self-report: Implications for research and practice*. Mahwah, NJ: Erlbaum.
- Suzuki, Y., Tamesue, D., Asahi, K., & Ishikawa, Y. (2015). Grit and work engagement: A cross-sectional study. *PLoS ONE*, *10*, e0137501.
- Tay, L., & Diener, E. (2011). Needs and subjective well-being around the world. *Journal of Personality and Social Psychology*, *101*, 354–365.
- Thalmayer, A. G., & Saucier, G. (2014). The questionnaire Big Six in 26 nations: Developing cross-culturally applicable Big Six, Big Five and Big Two inventories. *European Journal of Personality*, *28*, 482–496.
- Vainio, M. M., & Daukantaitė, D. (2016). Grit and different aspects of well-being: Direct and indirect relationships via sense of coherence and authenticity. *Journal of Happiness Studies*, *17*, 2119–2147.
- Vallerand, R. J., Blanchard, C., Mageau, G. A., Koestner, R., Ratelle, C., Léonard, M., ... Marsolais, J. (2003). Les passions de l'âme: On obsessive and harmonious passion. *Journal of Personality and Social Psychology*, *85*, 756–767.

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, 40, 472–481.

Von Culin, K. R., Tsukayama, E., & Duckworth, A. L. (2014). Unpacking grit: Motivational correlates of perseverance and passion for long-term goals. *Journal of Positive Psychology*, 9, 306–312.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Table S1

Table S2

Table S3

Table S4

Table S5

Table S6

Table S7

Table S8

Table S9

Figure S1

Figure S2

Figure S3

Figure S4

Figure S5

Figure S6

Figure S7

How to cite this article: Disabato DJ, Goodman FR, Kashdan TB. Is grit relevant to well-being and strengths? Evidence across the globe for separating perseverance of effort and consistency of interests. *Journal of Personality*. 2018;00:1–18. <https://doi.org/10.1111/jopy.12382>