Much Ado about Grit: A Meta-Analytic Synthesis of the Grit Literature

Marcus Credé
Department of Psychology, Iowa State University
mcrede@iastate.edu

Michael C. Tynan
Department of Psychology, Iowa State University
mtynan@iastate.edu

Peter D. Harms
School of Management, University of Alabama
pdharms@cba.ua.edu

ACCEPTED FOR PUBLICATION:
JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY
(Not the Paper of Record)

Corresponding Author: Marcus Credé, Department of Psychology, W112 Lagomarcino Hall, 901 Stange Road, Iowa State University, Ames, 50011-1041, mcrede@iastate.edu.
Abstract

Grit has been presented as a higher-order personality trait that is highly predictive of both success and performance and distinct from other traits such as conscientiousness. This paper provides a meta-analytic review of the grit literature with a particular focus on the structure of grit and the relation between grit and performance, retention, conscientiousness, cognitive ability, and demographic variables. Our results based on 584 effect sizes from 88 independent samples representing 66,807 individuals indicate that the higher-order structure of grit is not confirmed, that grit is only moderately correlated with performance and retention, and that grit is very strongly correlated with conscientiousness. We also find that the perseverance of effort facet has significantly stronger criterion validities than the consistency of interest facet and that perseverance of effort explains variance in academic performance even after controlling for conscientiousness. In aggregate our results suggest that interventions designed to enhance grit may only have weak effects on performance and success, that the construct validity of grit is in question, and that the primary utility of the grit construct may lie in the perseverance facet.

Keywords: grit; performance; meta-analysis; perseverance of effort; consistency of interest
Interest in non-cognitive variables as potential predictors and determinants of academic performance has been spurred by meta-analytic findings that variables such as study habits and study skills (Credé & Kuncel, 2008; Robbins et al., 2004), personality traits such as conscientiousness (Porapat, 2009, 2014), test anxiety (Hembree, 1988; Seipp, 1991), adjustment (Credé & Niehorster, 2012), emotional intelligence (Perera & DiGiacomo, 2013), and learning strategies (Credé & Phillips, 2011; Richardson, Abraham & Bond, 2012), exhibit relations with academic performance that are often comparable to relations between admissions test scores and academic performance (e.g., Hezlett et al., 2001). Many of these non-cognitive characteristics also appear to be responsive to interventions. For example, meta-analytic reviews have demonstrated that interventions can reduce anxiety (Hembree, 1988), and improve study skills (Hattie, Biggs, & Purdie, 1996) as well as social and personal skills (Durlak, Weissberg, & Pachan, 2010).

One additional non-cognitive variable that has received widespread attention and that has been widely touted as an important predictor of success and performance is a personality trait referred to as grit (Duckworth, Peterson, Matthews, & Kelly, 2007; Duckworth & Quinn, 2009). Grit is defined as “perseverance and passion for long-term goals” (Duckworth, Peterson, Matthews, & Kelly, 2007, p. 1087) and as “… not just resilience in the face of failure, but also having deep commitments that you remain loyal to over many years” (Duckworth as quoted in Perkins-Gough, 2013, p. 16). Duckworth et al. argue that grit can help to explain why some individuals perform better than their scores on ability tests might predict and that grit was a core contributor to the success of highly accomplished individuals such as Albert Einstein. Recently, Duckworth (2013a) has even argued that grit is as good or even a better predictor of success than cognitive ability; a strong claim given meta-analytic findings that cognitive ability correlates
about $\rho = .50$ with performance in academic and work settings (Sackett et al., 2012; Schmidt & Hunter, 1998).

Despite the widespread enthusiasm for grit as a potentially novel predictor and determinant of performance there are sound empirical and theoretical reasons why a critical reappraisal of the nature of the grit construct, its contribution to our understanding of performance, and its general position within the nomological network may be warranted. It is the goal of this paper to present findings from a meta-analytic synthesis of the rapidly growing empirical literature on grit in order to help shed light on the nature and construct validity of grit, and to highlight potentially new areas of inquiry for grit researchers. We begin by reviewing the literature relating to five core theoretical features of grit: 1) the proposed hierarchical structure of grit, 2) the relation of grit with performance, 3) the distinction of grit from conscientiousness, 4) the distinction of grit from cognitive ability and 5) the lack of group differences on grit.

**Grit as a Hierarchical Construct**

Grit is typically operationalized as a higher-order construct with two lower-order facets: “perseverance of effort” and “consistency of interest”. These two facets (hereafter referred to as perseverance and consistency), respectively refer to the tendency work hard even in the face of setbacks and the tendency to not frequently change goals and interests. Both are thought to contribute to success: persistence because the process of attaining mastery in a field often involves initial failures that the individual must persist through, and consistency because many hours of deliberate practice are normally required to achieve mastery (Ericsson, Krampe, & Tesch-Römer, 1993). That is, individuals who either disengage their efforts in the face of obstacles or who constantly change their interests are unlikely to ever engage in enough deliberate practice to achieve high levels of performance. The distinction between the two facets
is reflected in the subscales of the two primary self-report inventories used to measure grit: the Grit Scale (Duckworth et al., 2007) and the Short Grit Scale (Duckworth & Quinn, 2009) – both of which can be found on Angela Duckworth’s homepage (https://sites.sas.upenn.edu/duckworth/pages/research). Although some researchers examine perseverance and consistency as two separate constructs, most research on grit only reports findings at the level of the overall grit score.

This practice appears to be informed by two factors. First, Duckworth et al. (2007) in their discussion of the two facets state that “…neither factor was consistently more predictive of outcomes than the other, and in most cases, the two together were more predictive than either alone” (p. 1091). Second, Duckworth and Quinn (2009) reported examining the theoretical higher-order factor structure of grit using confirmatory factor analysis (CFA), and claimed support for the higher-order structure based on their finding that the higher-order model (comprised of two first-order factors and one second-order factor) exhibits significantly better fit than a single-factor model. However, the reported analysis is problematic because a model with one second-order factor and two first-order factors is not identified at the higher-order level (Kline, 2011); this means that no unique loadings onto the higher-order factor can be computed without additional constraints being imposed. However, it does not appear that Duckworth and Quinn imposed the type of equality constraints on the loadings of the first-order factors onto the higher-order factor that would be required to achieve identification because they report non-identical loadings of the first-order factors onto the second-order factor. Importantly, even if an equality constraint had been imposed at the higher-order level the resultant second-order model would have exhibited identical fit to a model with two correlated first-order factors and no second-order factor (see Credé & Harms, 2015 for a discussion). That is, a model in which the
persistence and consistency facets are simply two correlated constructs would exhibit identical fit to the higher-order model. Interestingly, Duckworth and colleagues had tested the fit of such a two-factor model in an earlier paper (Duckworth et al., 2007), and reported relatively poor fit for the model (i.e., CFI = .83, RMSEA = .11). The CFA strategy for determining whether a higher-order grit construct exists is therefore not particularly meaningful because standard indexes of model fit cannot be used to distinguish between a higher-order model and a model with two correlated factors.

A potentially more useful approach would involve examining the correlation between the two theoretical facets of grit; high correlations would suggest that a higher-order construct is plausible. However, empirical estimates of the strength of this relation exhibit substantial variability, with some reporting correlations that are close to zero (e.g., Chang, 2014; Datu, Valdez, & King; 2015; Jordan, Gabriel, Teasley, Walker, & Schraeder, 2015), while others have reported very strong correlations (e.g., Arslan, Akin, & Çitemel, 2013; Meriac, Slifka, & LaBat, 2015). A meta-analytic synthesis will help to establish a population estimate of the correlation between the two facets and thereby allow readers to make a more informed judgment about whether or not grit exhibits the hypothesized higher-order structure. This will in turn help to determine whether the practice of simply summing across the perseverance and persistence items to compute an overall grit score (as recommended by Duckworth et al., 2007) is reasonable, or whether the two facets should be considered separately.

**Grit as a Predictor of Success and Performance**

Proponents of grit as a predictor of performance have argued that between-person differences in grit can help to explain why two individuals with the same level of ability in a particular domain are often observed to perform at substantially different levels. Specifically,
individuals with high levels of grit are thought to be able to better utilize their capabilities because they are less distracted by short-term goals and less discouraged by the failures and setbacks that are commonly encountered in many performance domains. Indeed, Duckworth et al. (2007) described that the importance of grit for success had long been noted by prior research into highly accomplished individuals (e.g., Howe, 1999). Arguments for the importance of grit are also in line with work on the development of expertise that has highlighted the importance of sustained deliberative practice (e.g., Ericsson, Krampe, & Tesch-Römer, 1993; Krampe & Ericsson, 1996). Indeed, recent work by Duckworth, Kirby, Tsukayama, Berstein and Ericsson (2011) has explicitly tied grit to success in spelling bees via the mediating mechanism of deliberative practice. That is, individuals who are high on grit are more likely to engage in the amount of deliberative practice that is required to achieve expertise.

At the same time there are a number of theoretically plausible moderators of the grit-performance relation that suggest that the relation may not be strong in all - or even most - settings. First, the grit-performance relation may be moderated by the nature of the performance domain. Specifically, high levels of grit may be most useful when the task is difficult but well defined; that is, high levels of sustained effort and deliberative practice are required to succeed and the manner in which performance is to be attained is relatively clear (see MacNamara, Hambrick & Oswald, 2014). Thus, grit may be an excellent predictor of an individual’s ability to complete military basic training or succeed in well-defined academic tasks, but be less well related to performance on tasks that are very easy (thus not requiring grit) or performance on tasks that are novel and ill-defined and that therefore require both creativity and the willingness to abandon unsuccessful strategies (i.e., tasks on which grit may be counterproductive). Second, the grit-performance relation may be moderated by other individual differences such as ability
and meta-cognition (Credé & Phillips, 2011). That is, high levels of grit may not necessarily be adaptive unless it is accompanied by the general potential or ability to succeed in a domain and the ability to engage in the type of reflection and self-monitoring that the self-regulated learning literature (e.g., Zimmerman, 1990) and the social-cognitive view of learning (e.g., Ryan & Pintrich, 1997; Zimmerman, 1994, see also Kohn, 2014) has identified as important determinants of learning and performance. For example, cadets who lack some minimum level of physical ability to pass the highly strenuous tests of physical ability in a military academy are unlikely to benefit substantially from grit. Similarly, a cadet who is unable to recognize that a particular approach to studying for class material is not working is unlikely to perform well in academic courses.

Third, the grit-performance relation may be moderated by the level of grit itself. Very high levels of grit may become dysfunctional if they reduce the likelihood of help-seeking behaviors that have themselves been linked to performance (e.g., Karabenick, 2003) or if they increase the likelihood that an individual persists too long in attempting to solve a problem that is particularly difficult rather than spending their time on other, more solvable problems (see Lucas, Gratch, Cheng, & Marsella, 2015). This would, in turn, suggest that interventions designed to enhance grit levels may not benefit all individual equally. A recent large-scale evaluation of the impact of resiliency interventions (Paunesku et al., 2015) found that such interventions benefit primarily those students who were most at risk of dropping out of high schools and provided less benefit for other students. Meta-analytic evidence suggests similar effects for resilience interventions in working populations (Vanhove, Herian, Perez, Harms, & Lester, 2015). A similar non-linear relation may also exist for grit.
A more nuanced conceptualization of grit’s contribution to performance also seems warranted when reconsidering the biographical details of some of the highly accomplished scientific figures referenced by Duckworth et al. (2007) to highlight the importance of grit. For example, it is true that Albert Einstein persisted for many years in his attempt to develop the field equations that represent the mathematical description of General Theory of Relativity, but it is also true that Einstein persisted for years in pursuing an avenue of investigation that was based on an earlier mathematical error. Einstein also appears to have only resolved some of the challenging mathematical obstacles after consulting with the mathematicians Marcel Grossman and (later) David Hilbert (Earman & Glymour, 1978) who, according to some accounts (e.g., Parker, 2004), almost scooped Einstein in the development of mathematical framework for the General Theory of Relativity because of Einstein’s delay in seeking assistance. That is, persistence in this narrow case almost resulted in “failure” and Einstein only “succeeded” in this particular endeavor once he recognized his mathematical limitations and sought the help of more accomplished mathematicians.

The possibility that the grit-performance relation is not uniformly strong is also strongly suggested by an examination of the empirical literature. Some have found that grit scores are relatively strongly related to success as suggested by the initial findings by Duckworth and colleagues (e.g., Strayhorn, 2013), but many others (e.g., Chang, 2014; Cross, 2013; Davidson, 2014; Hogan, 2014; Sheehan, 2014) have failed to find strong relations between grit scores and indicators of success. This is particularly the case for studies examining academic success. Indeed, many of the reported relations between grit and academic success are weaker or equal to the relation of $\rho = .21$ and $\rho = .23$ that has been reported between conscientiousness and academic performance in two recent large-scale meta-analytic reviews (e.g., Porapat, 2009;
Richardson et al., 2013). Meta-analytic synthesis will not only help to clarify the strength of the relation between grit and success but will also help to clarify whether the observed variability in relations is simply a function of sampling error and other study artifacts (e.g., differences in the reliability in the measurement of variables), or if this variability reflects the presence of meaningful moderators.

A meta-analytic summary should also help to address possible confusion among readers of the grit literature about the ability of grit to predict the successful completion of rigorous programs. This confusion may have arisen because the authors of both of the foundational papers (Duckworth et al., 2007; Duckworth & Quinn, 2009) appear to confuse odds ratios with probabilities in their discussion of logistic regression results, resulting in incorrect inferences about the size of observed effects. This misunderstanding may have led readers to infer a much greater predictive power for grit scores than is warranted. For example, Duckworth and Quinn (2009) discussed the ability of grit scores to predict the successful completion a summer program for cadets from the United States Military Academy at West Point and interpret an odds ratio of 1.99 to mean that “Cadets who scored a standard deviation higher than average on the Grit-S were 99% more likely to complete summer training” (p. 171). This interpretation is incorrect because approximately 94% of all cadets successfully completed the program. A relatively small increase in the completion rate from, say, 95% to 97.5% associated with a one point increase in grit scores would, of course, represent an odds ratio of 2, but this is only a 2.6% increase in the likelihood of completing the program.

A meta-analytic synthesis will also help to establish whether either of the two facets of grit exhibit higher levels of criterion validity than the other or whether the two are largely equivalent in their relation with important outcomes as suggested by Duckworth and colleagues
(2007, p. 1091). This might, in turn, change the manner in which grit scores are presented and interpreted. Many grit researchers follow the recommendations by Duckworth et al. and examine only an overall grit score. However, facets are often better predictors than broad traits (e.g., Paunonen & Ashton, 2001), and substantial differences between the grit facets to predict important criteria might suggest that this scoring strategy should be revisited.

A meta-analytic synthesis of the strength of the relation between grit and success will also help to inform judgments about whether interventions designed to enhance grit are likely to have an impact on performance. Initial reports of the high predictive validity of grit scores and their relative independence from indicators of cognitive ability, combined with claims that grit can be taught (Perkins-Gough, 2013), has resulted in some schools implementing interventions designed to increase students’ levels of grit. For example, The Knowledge is Power Program (KIPP) network of public charter schools is training its teachers to foster grit in their pupils (Shechtman, DeBarger, Dornsife, Rosier, & Yarnall, 2013), while many school districts across the US are reportedly considering integrating the teaching of grit into curricula (Cohen, 2015). Grit was even highlighted as a promising focus of school interventions in a US Department of Education report (Shechtman et al., 2013). The time and resources that are likely to be devoted to grit-based interventions in schools are likely to be non-trivial and should therefore only be based on the best available knowledge about the role of grit in predicting and determining performance.

**The Distinction of Grit from Conscientiousness**

An interest in what Duckworth and colleagues refer to as grit, perseverance, and consistency is not new to psychology. Studies of attributes such as will power, tenacity, determination, persistence of motives, and volitional perseveration date back over 80 years (see Ryans, 1939 for an early review). More recently, researchers have investigated a variety of other
trait-like constructs that are characterized by persistence and consistency including proactivity (e.g., Crant, 1995), persistence (e.g., De Fruyt, Van de Wiele, & Van Heeringen, 2000), industriousness (e.g., Eisenberger, 1992; Jackson & Paunonen, & Tremblay, 2000), need for achievement (McClelland, 1985), conscientiousness, and some of the facets of conscientiousness such as industriousness, self-control, and order (Roberts, Chernyshenko, Stark & Goldberg, 2005). The conceptual similarities between these constructs and grit raises the possibility that the proponents of grit may have fallen victim to what Kelley (1927) referred to as the “jangle fallacy” – the belief that two things are different simply because they have different names. The contribution of grit to the psychological literature would, of course, be severely limited if the construct was simply a case of “old wine in new bottles” and it would therefore appear to be important to formally establish the discriminant validity of grit relative to these related constructs. There has however been almost no empirical investigation of the discriminant validity of grit from these other constructs using the types of methodologies (e.g., Multitrait-multimethod matrices, confirmatory factor analyses) commonly employed to determine discriminant validity, despite the fact that grit has been explicitly presented as a construct that is distinct from these previously examined constructs – particularly conscientiousness and need for achievement (Duckworth et al., 2007; Duckworth & Quinn, 2009; Perkins-Gough, 2013). The bivariate relation of grit with conscientiousness has often been reported by researchers but even for conscientiousness (and its facets) there are both empirical and theoretical reasons for suspecting that the overlap with grit may be stronger than is widely assumed.

For example, the definition of grit as “perseverance and passion for long-term goals” (Duckworth et al., 2007) is highly similar to the definitions given by Costa and McCrae (1992) for the self-discipline facet (“capacity to begin tasks and follow through to completion despite
boredom or distractions”) and the achievement striving facet of conscientiousness (“need for personal achievement and sense of direction”). This theoretical similarity is also reflected in the considerable similarity in the items that are found in the Duckworth et al. (2007) perseverance subscale and items in widely used inventories of conscientiousness such as those provided by the International Personality Item Pool (IPIP, Goldberg, 1999). For example, perseverance items such as “I finish whatever I begin” and “I am a hard worker” are very similar to IPIP items used to measure achievement striving such as “I carry out my plans” and “I work hard”. Items from the consistency scale overlap less strongly with items from conscientiousness inventories and are instead more similar to IPIP items used to measure the adventurousness facet of openness to experience and the IPIP measure of planfulness that is modeled after the Achievement via Conformance scale found in the California Personality Inventory (Gough, 1996).

A cursory examination of the empirical grit literature also suggests that the grit – conscientiousness relation may be much stronger than is commonly assumed. While some (e.g., Cooper, 2014) have presented evidence the grit is largely distinct from conscientiousness, numerous others have reported correlations between grit and conscientiousness that approach unity when correcting the observed correlations for unreliability. Reed, Pritschet and Cutton (2012), for example, report a correlation of $\rho = .92$ based on 1165 college students, Engel (2013) reports a correlation of $\rho = .95$ based on a smaller sample of 88, and Meriac, Slifka, and LaBat (2015) report a disattenuated correlation of $\rho = .98$ based on a sample of 322 students. Even Duckworth et al. (2007, 2009) report correlations between conscientiousness and grit scores that rise to $\rho = .97$ (N = 1,554); $\rho = .90$ (N = 706), $\rho = .83$ (N = 190), and $\rho = .80$ (N = 1,308) after correcting for the unreliability of both conscientiousness scores and grit scores. High correlations such as these have led some (e.g., MacCann & Roberts, 2010) to suggest that grit should be
considered a facet of conscientiousness – a position that seems theoretically plausible when considering that both grit and the self-control or self-discipline facet of conscientiousness focus on the deferment of short-term gain for long-term goals (see Costa & McCrae, 1992; Roberts, Bogg, Walton, Chernyshenko, & Stark, 2004). High observed correlations between grit and conscientiousness are also of concern when considering that a concurrent assessment of the same personality trait using different scales typically yields much lower correlations of around $r = .50$ (e.g., Pace & Brannick, 2010; Miller, Price & Campbell, 2012). Meta-analytic synthesis of the literature on the grit-conscientiousness relation will help clarify whether the strength of the relation is such that grit might be a case of the “old wine in new bottles” phenomenon.

**The Distinction of Grit from Cognitive Ability**

Grit is typically described as being largely distinct from cognitive ability (Duckworth et al., 2007; Duckworth & Quinn, 2009; Perkins-Gough, 2013), although Duckworth (2013a, 2013b) has also suggested a negative relation between grit and cognitive ability in noting that “… gritty people, on average, tend to be slightly less talented”. This distinction, if correct, suggests that grit might explain unique variance in performance over and above the substantial variance in performance accounted for by cognitive ability (e.g., Kuncel, Hezlett, Ones, 2004). Further, a finding that grit is largely orthogonal from general cognitive ability would also suggest that interventions designed to enhance grit levels might result in substantial increases in performance. Primary research findings have found broad support for the assertion that cognitive ability and grit are largely distinct. Many of the studies in this domain utilize admissions test scores as a proxy for cognitive ability test scores but this approach seems reasonable when considering that cognitive ability tests administered in research settings have questionable validity because many test takers will not be motivated to exert maximal effort on ability tests in
a low stake setting (Duckworth, Quinn, Lynam, Loeber, & Stouthamer-Loeber, 2011). Most commonly used college admissions test scores are highly correlated with cognitive ability test scores (e.g., Frey & Dotterman, 2004) and are also taken under maximal performance (i.e., high stakes) conditions and are therefore likely to represent a sound indicator of cognitive ability. Research examining the relation between grit scores and admissions test scores have largely found very weak relations (e.g., Chang, 2014; Duckworth et al., 2007; Eskreis-Winkler, Shulman, Duckworth, & Beal, 2014; Kelly, Matthews, & Bartone, 2014).

**Group Differences in Grit Scores**

Concerns about the reliance on cognitive ability tests for the prediction of success and performance have often revolved around the persistent finding that groups exhibit non-trivial mean score differences on such tests (e.g., Camara & Schmidt, 1999; Davis et al., 2013). A finding that grit exhibits smaller differences between groups is likely to make the construct more attractive in settings where scores are used for selection purposes or for making other high-stakes decisions because the likelihood of adverse impact on legally-protected groups is reduced. Prior research suggests only one type of group difference. Duckworth et al. suggested that grit might increase with age – a phenomenon that has also been observed for conscientiousness (Roberts, Walton, & Viechtbauer, 2006) – but prior findings from the personality literature (e.g., Costa, Terracciano, & McCrae, 2001; Foldes, Duehr, & Ones, 2008) suggest that differences across ethnicities and gender are likely to be more modest than those observed for ability measures. Empirical findings on the direction and strength of the relation between grit scores and demographic variables such as age (e.g., Engel, 2013, Eskreis-Winkler, Duckworth, Shulman & Beal., 2014), gender (e.g., Allen, 2014; Davidson, 2014), and ethnicity (e.g., Chang, 2014,
Eskreis-Winkler, 2014) have been mixed and a meta-analytic synthesis will help to clarify whether average grit scores are largely similar across groups.

**The General Position of Grit within the Nomological Network**

Grit has been not only been related to performance, cognitive ability, and conscientiousness but also to a wide array of other variables reflecting either states or traits. These include Big Five traits (e.g., Eskreis-Winkler et al., 2014), optimism (e.g., Lovering et al., 2015), psychological well-being (e.g., McCann & Roberts, 2010); suicide ideation (Blalock, Young, & Kleinman, 2015), intended persistence in academic programs (e.g., Bowman et al., 2015), and life satisfaction (e.g., Samson et al., 2011). A meta-analytic synthesis of the relation of grit with these other variables will further help to clarify the general position of grit within the broader nomological network.

**Interpreting Criterion-Related Validity Estimates**

Cohen’s (1988) guidelines for what constitutes small (r = .10), medium (r = .30), and large (r = .50) effect sizes are widely used to make describe the size of the relation between a predictor variable and a criterion variable. However, the American Psychological Association (Wilkinson & Task Force on Statistical Inference, 1999) has also encouraged researchers to place effect sizes in a practical and theoretical context. To this end we briefly discuss meta-analytic estimates of the criterion-related validity of various widely studied predictors of academic performance and retention. We use these meta-analytic estimates to inform our assessment of the relative ability of grit to predict academic performance and retention, although it is also important to note that even a relatively low criterion validity can be practically very important –
especially when the predictor provides information about the criterion that is not provided by other predictors and when the criterion is important.

Prior meta-analyses of predictors of academic performance have identified two variables that correlate at approximately $\rho = .50$ with academic performance in college: 1) indicators of cognitive ability such as scores on the SAT and 2) prior academic performance such as high school GPA (Sackett et al., 2012). Other predictors that correlate approximately at $\rho = .40$ with academic performance include study skills and study habits (Credé & Kuncel, 2008), 4) academic adjustment (Credé & Niehorster, 2012), academic self-efficacy (Robbins et al., 2004) and 6) class attendance (Credé et al., 2010). These appear to be the best known predictors of academic performance in college. Other variables that meta-analyses have shown to exhibit weaker but practically still very meaningful relations of around $\rho = .20$ with academic performance include: 1) specific learning strategies (Credé & Phillips, 2011), 2) emotional intelligence (Perera & DiGiacomo, 2013), 3) conscientiousness (Porapat, 2009), and 4) test anxiety (Hembree, 1988).

The ability to predict retention is typically weaker. Meta-analytic findings indicate that the best predictors are: academic self-efficacy ($\rho = .36$) and academic-related skills ($\rho = .37$, Robbins et al., 2004), institutional attachment ($\rho = .29$) and social adjustment ($\rho = .25$, Credé & Niehorster, 2012), high school grades ($\rho = .20$, Robbins, Allen, Casillas, Peterson, & Le, 2006), and SAT and ACT scores ($\rho = .17$, Mattern & Patterson, 2009; Robbins et al., 2006).

**Method**

**Search Strategy**
Potential sources for inclusion in our review were identified using keyword, abstract, and title searches of the PsycINFO, Dissertations Abstracts, and ERIC databases using the search term “grit”. This yielded a total of 778 potential data sources. These search results were supplemented by an examination of the reference lists of identified sources. We also examined the first 500 search results of the internet using the Google search engine and the search term “grit” to identify additional unpublished sources of information. Potential sources for inclusion were first screened by examining the abstract and title of the source and all possible sources were then examined more closely to determine if the reported data met the inclusion or exclusion criteria.

### Inclusion and Exclusion Criteria

Sources were included in our review if they reported on the Pearson correlation between scores on any of the Duckworth et al. (2007, 2009) measures of grit and other variables – or if they reported information that could be used to estimate the size of such a correlation (e.g., means and standard deviations for two criteria groups). The year of publication, source of the material, and country of origin of the data were not used to exclude any sources, although non-English sources were excluded. Sources were also excluded if they reported correlations for individuals below a middle school age because personality is still highly fluid at earlier ages, and because prior meta-analyses on personality as a predictor of achievement (Poraport, 2014) found that the strength of the relation at younger ages was very different to the relation at older ages. We also one study that only reported on significant correlations (and excluded non-significant correlations) because the inclusion of this data would have resulted in an upwardly biased effect size estimate. When studies did not report data in a format that could be coded and when these studies had been published in the last five years we attempted to contact the authors to request
the necessary information. Data from a total of 73 studies representing data from 88 unique samples and 66,807 individuals was ultimately included in the analyses.

**Coding Procedure**

All articles were coded by two of the authors using a systematic coding procedure, one of which has extensive experience coding articles for meta-analytic analyses. An accuracy check revealed 98.8% agreement in coding across the four most important coding categories. 60% of coding errors were errors of commission (e.g., incorrect coding of an effect size) and 40% were errors of omission (e.g., an effect size that could have been computed was not coded). All disagreements were resolved via discussion. Each correlation that was included in our review was described using ten coding categories: 1) the size of the correlation, 2) the sample size, 3) the reliability of the grit scores, 4) the reliability of the correlate scores, 5) the name of the correlate, 6) the source of grit ratings (self-ratings or other-ratings), 7) the source of the correlate data (self-ratings, other-ratings, records), 8) whether the grit scores reflected overall grit or either of the two facets: consistency and perseverance, 9) the source of the publication (peer-reviewed versus not peer-reviewed), and 10) the year of publication. Self-reported grades are very highly correlated with actual grades (Kuncel, Credé & Thomas, 2005) and we therefore included correlations with grades irrespective of whether the grade information was based on self-reports or were obtained from records. When sources reported correlations involving both self-reported grades and grades obtained from records we coded the correlations involving grades obtained from records. A summary of the coding of the most important variables is included in the Appendix.

**Transformations**
Our coding process involved three transformations of data. First, we used formulas presented by Hunter and Schmidt (2004) to calculate estimates of the correlation between grit and correlate variables when the original sources had artificially dichotomized the correlate variable (e.g., presenting grit scores for “low” and “high” scoring students). The artificial dichotomization of data results in downwardly biased estimates of the population correlation if such a correction is not made. Second, we used the formula presented by Ghiselli, Campbell and Zedeck (1981) to calculate composite correlations when the original source only presented correlations involving the facets of grit and/or facets of the correlate variable. For example, Bowman et al. (2015), presents correlations among the two facets of grit, Fall GPA, and Spring GPA; the six correlations among these four variables were used to arrive at an estimate of the correlation between overall grit and overall GPA. Mosier reliability estimates for composite variables (Mosier, 1943) were also calculated whenever possible. Finally, we computed point-biserial correlations between retention and grit when the mean and standard deviation of grit scores were reported for both the retained and non-retained group.

**Criterion Categories**

Grit researchers have examined the relation of grit with a wide variety of indicators of success. Meta-analytic synthesis requires a grouping of similar criteria with each other but because success criteria can be grouped in a wide variety of ways we present meta-analytic estimates for ten criterion categories. First, we present separate meta-analytic estimates of the relation between grit and high school GPA, college GPA, and post-graduate GPA. Second, we aggregate these into a broader general GPA criterion category (i.e., GPA across all three educational levels). For this general GPA criterion category we relied on the correlations for the more recent college GPA when authors reported correlations involving both high school GPA
and college GPA. We then also combined this general GPA category further with correlations involving grades in individual courses to form an Academic Performance criterion category. Third, we present meta-analytic estimates of the point-biserial correlation between grit and retention. Most studies examining retention do so in an academic or military setting (e.g., degree completion, completion of basic training) but one study also examined marital status as an indicator of retention (i.e., staying married versus getting a divorce or separating), and we therefore present meta-analytic estimates both with and without the study on marital retention. We also present findings for a criterion representing a collection of non-academic criteria comprised of performance in spelling bees, military settings, and athletics. Finally, we also present meta-analytic estimates of the relation between grit and the intent to persist in both college and with a particular employer.

**State, Trait, and Demographic Categories**

Grit researchers have examined the relation between grit and a variety of other variables that represent both relatively stable personality traits such as the Big Five personality traits and cognitive ability, and variables that have a strong mood and emotion component and that could therefore be described as falling somewhere along the state-trait continuum (e.g., happiness, depression, positive affect). We use the descriptors of these variables as given in our source articles to group these various state and trait variables into categories and computed meta-analytic estimates for those categories for which at least three effect sizes were reported. We also report meta-analytic estimates of the relation between grit and four demographic variables (gender, age, year in school and ethnicity).

**Statistical Method**
We used the Hunter and Schmidt (2004) interactive meta-analytic method based on a random-effects model to arrive at population estimates of the size of the relations between grit and other variables. The Schmidt and Le (2004) software was used to compute meta-analytic estimates of the relations involving grit and we corrected for unreliability in the measurement of the dependent variable and unreliability in the measurement of the independent variable. Grit scores are likely to be exhibit some level of range restriction in many samples but the absence of normative data on grit scores and variance in how grit is measured (e.g., number of items, number of response options) did not allow us to correct for range restriction.

**Corrections for unreliability.** In order to correct for the attenuating effect of measurement error on the size of the observed correlations we constructed reliability artifact distributions from the reliability information that was described in the included studies. These reliability distributions are described in Table 1. The included studies did not report information on the reliability of grades but in order to facilitate an apples-to-apples comparison with the recent meta-analytic summary of the relation between conscientiousness and academic performance by Porapat (2009) we corrected for the unreliability of grades using a distribution of reliability estimates for that was largely similar to the reliability estimates used by Porapat (2009). For GPA information taken from records we used the average of reliability estimates (alpha = .90) for all courses across four years as reported by Bacon and Bean (2006) while for self-reported GPA we used the operational validity estimates of .90 for college GPA and .82 for high school GPA as reported in the meta-analysis by Kuncel, Credé and Thomas (2005).

[INSERT TABLE 1 HERE]

**Other Analytic Decisions.** For the meta-analysis involving the retention criteria we took a dual analytical approach. More than half of the studies examining the relation between grit and
retention did not report means and standard deviations for those individuals who dropped out of a program and for those individual who stayed in the program. Instead these authors reported odds ratios, but odds ratios cannot be directly transformed into a point-biserial correlation when the independent variable is treated as a continuous variable. We therefore present two meta-analytic estimates for the grit-retention relation. We provide one estimate based purely on those studies that report data that could be transformed into a point-biserial correlation, and then provide another estimate that includes correlation values computed by taking the root of the Nagelkerke $R^2$ values reported for those studies that reported odds ratios from bivariate logistic regression models. Nagelkerke $R^2$ values tend to be too high as an estimate of the strength of the bivariate relation (Allison, 2014) but we include these values in order to provide readers with an estimate based on the most complete data. We remind readers that this estimate is likely to be upwardly biased.

We present meta-analytic estimates of the relations of overall grit (or the two grit facets) with other variables whenever at least three studies reported on such a relation. We summarize our meta-analytic findings for each relation using six pieces of information: 1) $k$ refers to the number of studies used to compute the estimates, 2) $N$ refers to the total sample size used to compute the estimate, 3) $r_{obs}$ refers to the sample-size weighted average observed correlation, 4) $\rho$ refers to the estimate of the population correlation, 5) $SD_\rho$ refers to the estimate of the standard deviation of effect sizes after taking into account the variability that is due to sampling error and differences in the reliability of measurement between studies, and 6) 10%CV and 90%CV represent the upper and lower bounds of the 80% credibility interval. The width of the credibility interval is indicative of the presence of undetected moderators. That is, wide credibility intervals indicate that the correlation can be expected to vary widely across settings.
In order to examine whether grit scores explain incremental variance in academic performance outcomes over and above the variance explained by conscientiousness we constructed a full meta-analytic intercorrelations matrix between grit, conscientiousness, and academic performance by importing the conscientiousness correlations of \( \rho = .21 \) for high school GPA and \( \rho = .23 \) for college GPA as reported by Porapat (2009), and using the average of these \( (\rho = .22) \) for overall academic performance. These correlation matrices were then used to perform hierarchical regression analyses based on the harmonic mean of sample sizes.

**Results**

Before proceeding with our primary analyses we first examined the data from the studies included in this meta-analytic review for publication and source bias.

**Publication and Source Bias**

We examine the possibility that the literature included in this meta-analytic review represents a biased sample of the research on grit in two ways. First, in order to examine whether the published and unpublished literature report grit-performance relations of different magnitudes we report separate meta-analytic estimates based on those studies that were published in peer-reviewed journals and all other studies (e.g., dissertations, conference presentations). Because of the limited number of total studies that report correlations at the facet level we only perform this analysis for overall grit. Results are presented in Table 2. In general the evidence for source bias is weak with only small differences in correlations reported for overall academic performance, the overall GPA criterion, and undergraduate GPA. None of the differences in correlations were significant at alpha = .05.

[INSERT TABLE 2 HERE]
Second, we use Egger’s Test of funnel-plot asymmetry (Eggers, Smith, Schneider, & Minder, 1997) to examine whether there is evidence in the meta-analysis for the exclusion of small studies with weak effects. Studies with small sample sizes that find weak effects may not be published and not found via literature searches and their exclusion may result in an overestimate of the strength of an effect. Egger’s Test regresses the standard normal deviate of the effect size for each sample onto the precision of the effect size estimate. The intercept of the regression line provides information about the size of any asymmetry; statistically significant negative intercepts suggesting that small studies with weak effects may have been suppressed from the literature. Because of the relatively small number of studies that examined any one relation we only perform a single test of asymmetry for the relation based on the largest number of studies: the relation between grit and overall academic performance. For this relation Egger’s Test indicated no significant asymmetry with the intercept being negative but weak and not significantly different from zero (a = -.15, p = .85).

The absence of evidence for strong source bias and publication bias suggests that our meta-analytic estimates are unlikely to be substantially biased in either a positive or negative direction by missing studies. We therefore present meta-analytic estimates of the relation of grit with criteria, state and trait variables, and demographic variables in Tables 3-5.

[INSERT TABLES 3 – 5 ABOUT HERE]

Relation between Perseverance and Consistency

Our meta-analytic estimate of the relation between perseverance and consistency (k = 17, N = 22,048, $\rho = .60$, $SD_\rho = .21$), indicates a generally strong relation although the width of the credibility interval suggests that the strength of this relation is substantially moderated. As an
exploratory follow-up analysis we compared the relation observed for the two different grit scales and found a stronger relation when researchers relied on the short grit scale (k = 11, N = 18,996, \( \rho = .66, \text{SD}_\rho = .15 \)) than when the original grit scale was used (k = 6, N = 3,052, \( \rho = .27, \text{SD}_\rho = .17 \)).

**Relations with Criteria**

Overall grit exhibits a relation with overall academic performance of \( \rho = .18 \) (k = 39, N = 13,141, \( \text{SD}_\rho = .11 \)) and \( \rho = .17 \) with the overall GPA criterion (k = 37, N = 12,601, \( \text{SD}_\rho = .10 \)). Among the academic performance criteria grit was approximately as strongly related to college GPA (k = 30, N = 10,526, \( \rho = .17, \text{SD}_\rho = .10 \)) as it was to high school GPA (k = 17, N = 6,364, \( \rho = .16, \text{SD}_\rho = .14 \)). Contrary to early assertions by Duckworth et al. (2007) that both facets predicted success outcomes equally well, the perseverance facet of grit exhibited much stronger relations with all academic performance criteria than the consistency facet. For example, perseverance correlated at \( \rho = .26 \) (k = 11, N = 5,221, \( \text{SD}_\rho = .12 \)) with overall academic performance while consistency correlated at only \( \rho = .10 \) (k = 11, N = 5,221, \( \text{SD}_\rho = .02 \)). A comparison of the correlations of perseverance and consistency with four academic performance criteria using the procedure for comparing correlated correlation coefficients described by Meng, Rosenthal, and Rubin (1992) showed that the correlations differed significantly (p<.001) in all four cases.

Grit correlated with retention at \( \rho = .12 \) when the marital success study is included (k = 11, N = 17,525, \( \text{SD}_\rho = .09 \)), at \( \rho = .18 \) (k = 10, N = 11,163 \( \text{SD}_\rho = .03 \)) when it was excluded, and at, \( \rho = .16 \) (k = 5, N = 2,705, \( \text{SD}_\rho = .06 \)) if the upwardly biased correlations estimated from Nagelkerke R^2 values are excluded. Grit was correlated at \( \rho = .21 \) (k = 7, N = 4,116, \( \text{SD}_\rho = .00 \)) with performance in non-academic domains. The relation between grit and the intent to persist in
college and with the current employer was $\rho = .18$ (k = 5, N = 3,967, SD$\rho$ = .00), and $\rho = .15$ (k = 4, N = 519, SD$\rho$ = .00) respectively.

**Relation with State and Trait Variables**

Consistent with the claim that grit and cognitive ability are largely orthogonal, grit exhibited only a very weak relation with cognitive ability (k = 21, N = 11,513, $\rho = .05$, SD$\rho$ = .12). Similarly weak relations with cognitive ability were also observed for both the perseverance facet (k = 5, N = 2,204, $\rho = -.01$, SD$\rho$ = .04) and consistency facet (k = 5, N = 2,204, $\rho = .00$, SD$\rho$ = .00).

Grit exhibited much stronger relations with other trait variables. Conscientiousness was very strongly correlated with overall grit (k = 22, N = 18,826, $\rho = .84$, SD$\rho$ = .07) and also with perseverance (k = 8, N = 4,967, $\rho = .83$, SD$\rho$ = .14) and consistency (k = 8, N = 4,967, $\rho = .61$, SD$\rho$ = .17). Grit also exhibited a very strong relation with self-control (k = 4, N = 2,615, $\rho = .72$, SD$\rho$ = .05), a variable that is often seen to be a facet of conscientiousness (e.g., Roberts et al., 2005). Grit also exhibited a relatively strong relation with emotional stability (k = 14, N = 14,501, $\rho = .41$, SD$\rho$ = .04) but this should not be surprising when considering that low levels of emotional stability are likely to be associated with an inconsistency of interest because negative affect states may be interpreted as a signal that the activity being engaged in is no longer interesting. Grit also exhibited relatively strong relations with a number of other variables that are sometimes presented as having a causal influence on success and performance, including: self-control (k = 4, N = 2,615, $\rho = .72$, SD$\rho$ = .05), generalized self-efficacy (k = 3, N = 1,908, $\rho = .43$, SD$\rho$ = .11), mental toughness (k = 6, N = 3,817, $\rho = .46$, SD$\rho$ = .08), positive affect (k = 3, N = 670, $\rho = .46$, SD$\rho$ = .03), and depression (k = 5, N = 3,865, $\rho = -.48$, SD$\rho$ = .12).
Relation with Demographic Variables

As expected, the relations between grit and demographic variables such as gender (k = 25, N = 18,750, \( \rho = .05, \text{SD}_\rho = .07 \)), year in school (k = 4, N = 2,961, \( \rho = .05, \text{SD}_\rho = .05 \)), and ethnic minority status (k = 9, N = 15,261, \( \rho = .01, \text{SD}_\rho = .01 \)) were all very weak with the exception of age which exhibited a slight positive correlation with overall grit (k = 22, N = 12,349, \( \rho = .12, \text{SD}_\rho = .04 \)) in line with the prediction by Duckworth et al. (2007) that grit would increase with age. This increase is similar to the general increase in conscientiousness observed with age (Roberts et al., 2006).

[INSERT TABLE 6 HERE]

Incremental Validity

The incremental validity estimates from these regression results are summarized in Table 6. Results for Model 1 indicate that overall grit explains no variance in either overall academic performance or high school GPA after controlling for conscientiousness, and explains only a very small amount of incremental variance in college GPA (i.e., \( \Delta R = .004 \)). Importantly, conscientiousness explains incremental variance in these outcomes if first controlling for overall grit. Results for Model 2 show that perseverance explained a substantial amount of incremental variance in overall academic performance (\( \Delta R = .040 \)), high school GPA (\( \Delta R = .085 \)), and a somewhat lower amount for college GPA (\( \Delta R = .023 \)). Consistency explained almost no unique variance in the three criteria after controlling for either conscientiousness (see Model 3) or both conscientiousness and persistence (Model 4) and the negative sign of the regression coefficients for consistency for the overall academic performance and college GPA criteria also suggests a possible suppressor effect. Perseverance explained relatively large amounts of unique variance in
three criteria even after controlling for both conscientiousness and consistency (Model 5).
Overall, the incremental validity findings suggest that both overall grit and the consistency facet of grit add little to our ability to understand or predict academic performance, while the perseverance facet does offer an important improvement in explanatory power.

**Discussion**

Proponents of grit have asserted that grit is a higher-order construct composed of a perseverance facet and a consistency facet, that grit scores are highly predictive of success (and a better predictor of success than cognitive ability), and that grit scores provide information about individuals that is meaningfully distinct from conscientiousness. Three primary findings from our meta-analytic review of the grit literature suggest that the validity of these assertions may need to be revisited; although our findings also suggest that a revised approach to the study of grit may still hold value for our understanding of the determinants of performance.

First, our findings indicate that the current evidence does not support the claim that grit is a higher-order construct that is characterized by two lower-order facets. The original factor analytic studies could not speak to the presence of a higher-order factor structure because of methodological limitations, and our results indicate that the practice of combining perseverance scores and consistency scores into an overall grit score appears to result in a significant loss in the ability to predict performance. That is, perseverance is a much better predictor of performance than either consistency or overall grit and should therefore probably be treated as a construct that is largely distinct from consistency in order to maximize its utility.

Second, overall grit exhibits relations with academic performance and retention that are only modest and that do not compare favorably with other well-known predictors of academic
performance such as cognitive ability (Sackett et al., 2012), study habits and skills (Crede & Kuncel, 2008), and academic adjustment (Crede & Niehorster, 2012). Indeed, meta-analytic reviews of the literatures for some of these other predictors report correlations with academic performance and retention that are more than twice as big as those observed for overall grit in this review. At the same time it should be remembered that variables that exhibit small to moderate effect sizes can still be very useful in high-stakes settings because even marginal improvements in individuals’ performance - or organizations ability to predict this performance – can have very meaningful positive effects. For example, a grit intervention that increasing the retention rate in college by even a single percentage point would potentially benefit thousands of college students. Similarly, even a small increase in the ability to predict future performance in a selection setting may yield very substantial financial benefits for an organization (see Hunter & Hunter, 1984 for a discussion). Such a benefit would be particularly large if the variable in question reflected information about individuals that was distinct from the information reflected by other well-known predictors of performance and retention.

Our third primary finding suggests that the incremental value of grit for the prediction of performance is likely to be limited. Grit scores exhibited very strong correlations with conscientiousness and with self-control – a facet of conscientiousness. Indeed, the size of the correlation ($\rho = .84$) with overall conscientiousness is so strong as to not only limit the incremental value of grit scores for the prediction of performance over and above conscientiousness but also suggest that grit may be redundant with conscientiousness. Indeed, the correlation between overall grit and conscientiousness, and between persistence and conscientiousness ($\rho = .89$) is much stronger than what is typically found between scores on two different global measures of conscientiousness ($\rho = .63$; Pace & Brannick, 2010). This, in turn,
suggests that grit research may have fallen victim to the jangle fallacy and that grit as currently measured is simply a repackaging of conscientiousness or one of the facets of conscientiousness. McCornack (1956), of course, illustrated that two variables can be very strongly correlated but still exhibit very different correlations with a third variable but the meta-analytic estimates of the relation between overall grit and GPA in middle/high school ($\rho = .16$) and college ($\rho = .17$) are largely identical (if somewhat weaker) to those reported for conscientiousness in the recent meta-analytic review by Porapat (2009): $\rho = .21$ for middle/high school GPA and $\rho = .23$ for college GPA.

Although our findings indicate that a critical re-appraisal of the grit construct is warranted, three meta-analytic findings reported in this paper do hold some promise for proponents of grit as a predictor of success and as a potential focus of interventions. First, grit predicts retention approximately as well as many more traditional predictors of retention such as cognitive ability and high school grades – although not as well as some other non-cognitive predictors. This suggests that the assessment of grit may be useful in settings in which retention is problematic (e.g., higher education) because it may allow researchers to identify individuals who might benefit the most from interventions that target grit or offer assistance in some other fashion. Second - as noted earlier - our meta-analytic results show that the perseverance of effort facet of grit exhibits substantially higher criterion validity for the grade criteria than the consistency of interest facet. Indeed, the observed criterion validity of perseverance for the high school GPA criterion is also significantly higher than the criterion validity observed for overall grit scores and also for conscientiousness (Poropat, 2009). This suggests that the focus of the grit researchers should shift to perseverance as the most promising avenue of future research. Third, our hierarchical regression results based on meta-analytic estimates suggest that perseverance of
effort scores explain incremental variance over and above conscientiousness in the various grade criteria. This is, of course, encouraging, but the only moderately high correlations among scores on most personality traits assessed via two different inventories (see Pace & Brannick, 2010) mean that such incremental validity findings would be observed even if grit was simply a different manifestation of conscientiousness as our other results suggest.

Even modest predictor-criterion relations can be very important in applied settings, especially when individuals’ standing on the predictor can be impacted by simple interventions. Whether it is possible to enhance grit via interventions is not yet clear although evidence that social and personal skills as well as resiliency are responsive to interventions (Durlak et al., 2010; Paunesku et al., 2015) suggest that grit interventions may have some positive effect. Although we do believe that our results regarding the validity of the perseverance facet offer some promise we also believe that our overall results should lead to a re-evaluation of the appropriateness of planned or existing grit interventions. Schools and colleges have limited resources to devote to interventions and are likely to be best served by focusing those resources on variables that have been demonstrated to be 1) most strongly related to performance and persistence/retention and 2) responsive to interventions. Fortunately there are a number of variables that meet both of those requirements. For example, study skills and habits have been shown to correlate approximately $\rho = .40$ with college GPA (Credé & Kuncel, 2008), while Hattie et al. (1996) showed that study skills interventions can have moderate positive effects on study skills. College students’ adjustment to college has been shown to be similarly predictive of academic performance ($\rho = .39$ for academic adjustment), is also one of the best predictors of retention in college ($\rho = .29$ for institutional attachment) and can be slightly improved by simple interventions such as orientation programs (see Credé & Niehorster, 2012 for a review). Simple
class attendance is also very strongly related to academic performance, and making class attendance compulsory appears to dramatically reduce the proportion of students who fail a class (Credé et al., 2010). Study skills and study habits, adjustment to college, and class attendance are thus far more strongly related to academic performance and retention than grit, and there is sound evidence that interventions can improve students’ standing on these constructs (especially for study skills and habits).

Limitations and Future Research

Meta-analyses are limited by the nature and quality of the data present in a literature. As such this meta-analysis of the grit literature has some notable limitations. First, the literature relating grit to academic performance is primarily based on concurrent designs. This, in turn, may have resulted in inflated estimates of the grit-academic performance relation because individuals’ knowledge of their academic performance may influence their responses to the measure of grit. Second, there are at least three reasons why the grit literature may be characterized by a non-trivial amount of range restriction. All of the examined studies relied on self-reports of grit and the social desirability of grit items may have resulted in range restriction in self-reported grit scores. Individuals may also generally not be aware of their true level of grit and unintentionally report inflated levels of grit (Kruger & Dunning, 1999). Further, samples drawn from populations that have been selected based on prior performance may exhibit some range restriction on grit. For example, cadets at the US Military Academy at West Point have likely exhibited outstanding academic performance in high school and may therefore have a lower range of grit scores than the range found in the general population. We were unable to correct for range restriction in our meta-analysis because of the lack of normative data and the variability in how the grit scales were used by researchers but future research may be able to
estimate the level of range restriction that is present in samples. Third, many of the studies examining the relation between grit and retention were characterized by very high base rates of retention (i.e., low rates of dropout). Duckworth et al. (2007) for example report data on one sample from the United States Military Academy in which 94.2% of the sample were retained through the examined period. Such low base rates severely attenuate the size of the correlation that can be observed. In such circumstances meta-analyses could make corrections for range restriction, but such a correction would require information about the size of the standard deviation for the retention criteria in the general population and we are not aware of a reasonable estimate for this value. Finally, although the empirical grit literature is sufficiently large to allow us to comment with relative confidence on average population effect sizes the literature is not yet large enough to allow moderator analyses characterized by high power (Hunter & Schmidt, 2004). An exploration of the reasons for the occasionally wide credibility intervals will require the accumulation of further data.

We believe that future research in this domain should consider five broad issues. First, researchers should attempt to examine whether grit exhibit the type of stability that is associated with other personality traits or whether it is responsive to interventions. Grit interventions will need to be tested to evaluate the malleability of grit but there are sound theoretical reasons why such interventions may be effective. Eisenberger (1992) argued that industriousness – a construct that is similar to grit – can be acquired via reinforcement and that repeated reinforcement for high effort (i.e., grit) can eventually result in a generalized increase in effort across tasks even when these tasks are not extrinsically reinforced. This work will not only require long-term experimental manipulations in the form of reinforcements for high effort but should help to establish the degree to which grit is truly trait-like and also help to clarify the type of grit.
interventions that are likely to be most effective. Second, grit researchers should consider examining criteria that span to different domains (e.g., work settings), a greater range of difficulty and a greater variety of task types (e.g., intellective tasks versus creative tasks). This may help to establish boundary conditions for the influence of grit on success and performance. Third, grit researchers should consider examining the potential moderators of the grit-performance relation discussed earlier: the moderating role of the performance domain; the moderating role of individual differences such as ability and meta-cognition; and the moderating role of the level of grit. Fourth, it may also be useful to examine the degree to which scores on measures of grit are related to scores on measures of motivation. A popular definition of motivation is that it reflects “an individual’s intensity, direction, and persistence of effort toward achieving a goal” (e.g., Robbins, Judge, & Campbell, 2010) and as such bears clear conceptual similarities to grit. Finally, it is possible that the grit literature may benefit from a refinement of the grit scale using method based on Item Response Theory. It is unlikely that the relatively short measures of grit are equally good at assessing low, medium, and high levels of grit. This lack of depth and breadth in item content could lead to attenuation of the reported effects (see Credé, Harms, Niehorster, & Gaye-Valentine, 2012). Better measures of grit would not only help to clarify the nature of the grit-performance relation but would also be important for the evaluation of future grit interventions.

Conclusion

Grit as a predictor of performance and success and as a focus of interventions holds much intuitive appeal, but grit as it is currently measured does not appear to be particularly predictive of success and performance and also does not appear to be all that different to conscientiousness. We hope that greater rigor in scale development, a greater focus on the perseverance facet, and a
more nuanced approach in study design will help future grit researchers to develop boundary conditions for grit in its role as in influence on performance and success.
References
(* denotes inclusion in meta-analysis)


doi:10.1080/20445911.2014.995104


doi:10.1037/0022-3514.92.6.1087


Table 1

Artifact Distributions used for Meta-Analytic Computations

<table>
<thead>
<tr>
<th>Variable</th>
<th>$k_\alpha$</th>
<th>Mean $\alpha$</th>
<th>SD $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Grit</td>
<td>46</td>
<td>.79</td>
<td>0.07</td>
</tr>
<tr>
<td>Perseverance</td>
<td>10</td>
<td>.71</td>
<td>0.13</td>
</tr>
<tr>
<td>Consistency</td>
<td>11</td>
<td>.74</td>
<td>0.11</td>
</tr>
<tr>
<td>Overall Academic Performance</td>
<td>4</td>
<td>.88</td>
<td>0.04</td>
</tr>
<tr>
<td>Overall GPA</td>
<td>4</td>
<td>.88</td>
<td>0.04</td>
</tr>
<tr>
<td>College GPA</td>
<td>2</td>
<td>.90</td>
<td>0.00</td>
</tr>
<tr>
<td>Graduate School GPA</td>
<td>2</td>
<td>.90</td>
<td>0.00</td>
</tr>
<tr>
<td>High School GPA</td>
<td>2</td>
<td>.86</td>
<td>0.06</td>
</tr>
<tr>
<td>Intent to Persist at Current Employer</td>
<td>4</td>
<td>.93</td>
<td>0.05</td>
</tr>
<tr>
<td>Intent to Persist in College</td>
<td>2</td>
<td>.70</td>
<td>0.25</td>
</tr>
<tr>
<td>Cognitive Ability</td>
<td>2</td>
<td>.87</td>
<td>0.07</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>11</td>
<td>.75</td>
<td>0.14</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>17</td>
<td>.79</td>
<td>0.09</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>11</td>
<td>.81</td>
<td>0.07</td>
</tr>
<tr>
<td>Extroversion</td>
<td>11</td>
<td>.83</td>
<td>0.10</td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>12</td>
<td>.76</td>
<td>0.06</td>
</tr>
<tr>
<td>Generalized Self-Efficacy</td>
<td>2</td>
<td>.90</td>
<td>0.06</td>
</tr>
<tr>
<td>Optimism</td>
<td>1</td>
<td>.86</td>
<td></td>
</tr>
<tr>
<td>Gratitude</td>
<td>3</td>
<td>.81</td>
<td>0.08</td>
</tr>
<tr>
<td>Mental Toughness</td>
<td>3</td>
<td>.84</td>
<td>0.12</td>
</tr>
<tr>
<td>Hope</td>
<td>4</td>
<td>.77</td>
<td>0.17</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>1</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td>Life Satisfaction</td>
<td>5</td>
<td>.87</td>
<td>0.03</td>
</tr>
<tr>
<td>Work Satisfaction (College and Job)</td>
<td>1</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>5</td>
<td>.86</td>
<td>0.05</td>
</tr>
<tr>
<td>Happiness</td>
<td>3</td>
<td>.69</td>
<td>0.14</td>
</tr>
<tr>
<td>Resiliency</td>
<td>3</td>
<td>.88</td>
<td>0.07</td>
</tr>
<tr>
<td>Self-Control</td>
<td>3</td>
<td>.84</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note: $k_\alpha$ is the number of reliability estimates in distribution, Mean $\alpha$ is the mean of the reliability estimates, SD $\alpha$ is the standard deviation of the reliability estimates.