Other-rated personality and academic performance: Evidence and implications.

Abstract

Considerable gaps remain in teachers' and students' understanding of factors contributing to learning and educational outcomes, including personality. Consequently, current knowledge about personality within educational settings was reviewed, especially its relationships with learning activities and academic performance. Personality dimensions have previously been shown to be related to learning strategies and activities, and to be reliably correlated with academic performance. However, personality is typically self-rated, introducing methodological disadvantages associated with informational and social desirability biases. A meta-analysis of other-rated personality demonstrated substantially higher correlations of academic performance with all of the dimensions of the Five-Factor Model of personality, which were not accounted for by associations with intelligence. The combined association of academic performance with all of the Five-Factor Model dimensions was one of the largest so far reported in education. The findings have implications for personality measurement. Teachers are able to assess students' personalities to match educational activities to student dispositions, while students' development of learning capacities can be facilitated by feedback on how their personalities are linked with effective learning.

Running Head: Personality & educational outcomes

Keywords: Other-ratings; Personality; Five-Factor Model; Academic Performance; Meta-analysis

1. Introduction

From ancient schooling and examinations in China (Bowman, 1989) and Egypt (David, 1998) to modern educational systems, learning and academic performance have long been centrally-important outcomes for individuals and their social and economic advancement. Within advanced contemporary economies, most people undergo formal education for a quarter of their expected life-spans, an effort that accounts for direct expenditure equivalent to an average of 6.2% of national economic output (OECD, 2007). This is not surprising, given the substantial multiplier effects associated with educational investments, both for economies as a whole and for individual career and economic success (Strenze, 2007).

Yet regardless of the importance of learning and subsequent academic performance, people are not good at recognizing what actually helps them to learn, adopting practices that are somehow ill-informed by their own experience (Bjork, Dunlosky, & Kornell, 2013). Worse still, “Nor do customs and standard practices in training and education seem to be informed, at least reliably, by any such understanding” (Bjork et al., 2013, p.419). This situation is at least partly due to the fact that “… despite all the pronouncements, there is not a detailed understanding of the mechanisms by which students turn their learning experiences into knowledge” (Lee & Anderson, 2013, p.463). There is still much to learn about learning.

One of the areas in which both educators and learners have been under-informed is the role of individual differences in learning and education, especially with respect to temperament and personality. Late last century, Rothbart and Jones (1998) argued that teachers need to be able to recognize differences in student temperament (and by inference, personality) in order assist different students to learn how to learn in ways that match their particularities. This point was strengthened by more recent meta-analyses by Poropat (2009) and Richardson, Abraham, and Bond (2012), demonstrating reliable associations of

personality with learning and academic performance. Yet in an otherwise thorough review of the key factors that contribute to learning, Dunlosky, Rawson, Marsh, Nathan, and Willingham (2013) overlooked temperament and personality in their consideration of factors that can affect learning outcomes.

In part, this oversight makes sense. Dunlosky et al. (2013) were interested in identifying "easy-to-use learning techniques that could help students achieve their learning goals" (p.4), yet personality assessment is typically less than easy-to-use, with little clarity about how knowledge of personality may inform educators. A range of problems exist with many methods of assessing personality, typically limiting both validity of measurement and observed associations with educational outcomes. Consequently, in this report I provide a brief review of theory and research on personality in education, leading to a meta-analytic examination of an alternative approach to personality assessment that has potential to be both easier for teachers to apply and more informative for students.

1.1 Individual Differences in Education

Considerable research efforts have attempted to document individual difference contributors to educational success (Kuncel & Hezlett, 2007), which was a major catalyst for early research on mental abilities, prompting the work that led to the identification of the general intelligence factor or *g* (Spearman, 1904, 1927). Academic performance and *g* have been so closely associated empirically and conceptually that the former is often the primary validating criterion for tests of the latter (Chamorro-Premuzic & Furnham, 2006).

In contrast, currently-dominant measures of personality were developed without educational outcomes in mind. Instead, the FFM was based upon the lexical hypothesis, the idea that the most important dimensions of personality are reflected by the largest number of descriptors within natural languages (Allport & Odbert, 1936; Saucier & Goldberg, 1996).

Lexical-hypothesis-inspired factor-analyses of personality descriptors (e.g., Digman, 1989; Tupes & Christal, 1961) have produced a range of personality models, most notably the Five-Factor Model (FFM), composed of the following dimensions: Agreeableness (reflecting tendencies to be peaceful, tolerant, warm and accommodating), Conscientiousness (encompassing diligent, dutiful and hard-working), Emotional Stability (self-assured and imperturbable versus the opposite pole, often referred to as Neuroticism, covering anxious, vulnerable and fearful), Extraversion (including sociable, talkative and lively), and Openness (creative, smart, intellectual and curious) (Saucier, 2009). The popularity of the FFM has enabled a series of meta-analyses on the relationship between personality and performance, establishing that FFM factors are reliably associated with academic performance, and that this association is independent of *g* (Connelly & Ones, 2010; Poropat, 2009; Richardson et al., 2012). In the following sections, theoretical and empirical reasons for the associations of the FFM dimensions with learning and academic performance are briefly summarized.

1.1.1 Agreeableness.

The personality factor of Agreeableness is linked with prosocial tendencies, in contrast with antagonistic behavior (Shiner & Caspi, 2003). A range of reasons have been proposed as to why Agreeableness may be associated with academic performance, including its association with positive relationships, which should facilitate learning (Saklofske, Austin, Mastoras, Beaton, & Osborne, 2012), and the extent to which more agreeable students are extrinsically motivated (Clark & Schroth, 2010), resulting in greater compliance with educational instruction. Consistent with this, students who are higher on Agreeableness have greater compliance with homework directions (Lubbers, Van der Werf, Kuyper, & Hendriks, 2010), and better time management and effort regulation (Bidjerano & Dai, 2007). Yet this compliance appears more social than educational in focus: meta-analytic correlations of

Agreeableness with academic performance at secondary and tertiary level have been modest (.05 & .06: Poropat, 2009; .07: Richardson et al., 2012).

1.1.2 Conscientiousness.

Conscientiousness is arguably the oldest of modern personality dimensions, with Webb (1915) reporting a factor similar to Conscientiousness that was subsequently confirmed by Deary (1996). Conscientiousness is also the factor with the strongest and most reliable association with academic performance at secondary and tertiary levels (.21 & .23: Poropat, 2009; .23: Richardson et al., 2012) a level comparable with that reported for correlations of measures of intelligence with grade-point-average (GPA) (.24 & .23: Poropat, 2009; .21: Richardson et al., 2012).

Although, Conscientiousness is typically associated with behavior rather than emotion (Pytlik Zillig, Hemenover, & Dienstbier, 2002), it provides significant advantages in stressful situations, with more conscientious students being more likely to focus on learning tasks (Saklofske et al., 2012), and to use problem-focused coping (MacCann, Lipnevich, Burrus, & Roberts, 2012). These students consequently are less tense and more confident (Cianci, Klein, & Seijts, 2010; MacCann, Lipnevich, et al., 2012), which in turn should allow them to stay focused on educational activities leading to greater learning.

More importantly, Conscientiousness has been both theoretically and empirically-linked with effortful control (De Pauw, Mervielde, & Van Leeuwen, 2009; Rothbart, 2007), and self-control (MacCann, Duckworth, & Roberts, 2009; Moffitt et al., 2011), the capacity to regulate attention in order to control emotion and action (Duckworth, Quinn, & Tsukayama, 2012). As a consequence, students who are higher on Conscientiousness make better use of self-regulation strategies such as time-management (Bidjerano & Dai, 2007), focusing on goals, making plans, and complying with rules (Roberts, Jackson, Fayard, Edmonds, &

Meints, 2009), as part of a strategic learning approach in which they adapt their activities to educational demands (Swanberg & Martinsen, 2010).

This association with strategic learning is echoed in Richardson et al.'s (2012) suggestion that effort regulation mediates the relationship between Conscientiousness and academic performance. If so, this is an important issue for educators because of the central role of self-regulation within learning (Bjork et al., 2013). For example, scripts and rubrics appear to contribute to learning largely because of their effect on self-regulation (Panadero, Tapia, & Huertas, 2012). One specific example of this was provided by MacCann, Fogarty, and Roberts (2012), who found that the effort regulation activity of time management mediated the Conscientiousness-academic performance, apparently because conscientious students allocate their study time more effectively. Likewise, conscientious students engaging in more learning-related activities (Steinmayr, Bipp, & Spinath, 2011), leading to better learning outcomes (Zhang, Nurmi, Kiuru, Lerkkanen, & Aunola, 2011). Consequently, the ability to accurately assess Conscientiousness should provide a basis for educators to tailor learning activities in a manner supportive of individual students' strengths and weaknesses.

1.1.3 Emotional Stability.

Emotional Stability, the opposite pole from neuroticism, is negatively correlated with negative affectivity and reaction intensity, and positively correlated with distractibility control (De Pauw et al., 2009). Consequently, when less emotionally-stable students are faced with anxiety-provoking conditions, such as impending assessments, they tend to ruminate on negative thoughts and feelings, distracting students from learning tasks and hence reducing associated learning (Kircanski, Craske, & Bjork, 2008). Such students focus on worrying about errors rather than upon the errors themselves, which impedes learning from those errors (Zhao, 2011), while emotionally stable students are more skilled at staying focused on learning activities (Lubbers et al., 2010). When less emotionally- stable students get

distracted or avoid learning situations, their capacity to engage with challenging learning situations is limited, reducing their exposure to the ‘desirable difficulties’ that contribute to effective learning (Graesser, 2009). Yet self-rated Emotional Stability has a reliable correlation with academic performance at primary level (.20: Poropat, 2009), there are non-significant correlations at secondary and tertiary levels of education (.01 & -.01: Poropat, 2009; -.01: Richardson et al., 2012).

1.1.4 Extraversion.

Extraversion is one of the key components of interpersonal behavior (Saucier, Ostendorf, & Peabody, 2001). In one of the first large-scale studies of personality in education, Eysenck and Cookson (1969) found that extraversion was correlated with academic performance, which they argued was due to extraverted students interacting more with their teachers, leading to greater learning and higher academic performance. This correlation of extraversion with academic performance was later confirmed by Poropat (2009), but only for students in primary education. At higher levels of education, correlations of academic performance with self-rated Extraversion are small enough to be meaningless in a practical sense (-.03 & -.01: Poropat, 2009; -.03: Richardson et al., 2012). So, although more extraverted students may get greater attention leading to higher performance at primary level, the reduced strength of teacher-student relationships at higher levels of education appear to eliminate this effect.

1.1.5 Openness.

Of all of the FFM dimensions, Openness is the most difficult to clearly identify (DeYoung, Peterson, & Higgins, 2005), both because of the difficulty of replicating the factor across cultures (Saucier & Goldberg, 2001) and its complex nature. Openness is commonly seen as comprised of two facets, one of which is more focused upon aesthetics while the other reflects thinking and curiosity (Von Stumm, Hell, & Chamorro-Premuzic, 2011). It is

this second component, sometimes labeled Intellect, which is both most associated with intelligence and appears to be most closely tied to learning and academic performance (Mussel, 2013). One explanation for this association is that Openness-related willingness-to-learn affects the level of intellectual investment in learning, and subsequent academic performance (Ziegler, Danay, Heene, Asendorpf, & Buhner, 2012), while different facets of Intellect are meaningfully related to different components of academic performance (Mussel, 2013).

After Conscientiousness, Openness is the personality dimension with the strongest correlation with academic performance (.12 & .07: Poropat, 2009; .09: Richardson et al., 2012). The thinking and curiosity aspects of Openness (Bernard, 2010) is expressed in a deep approach to learning, in which students follow their intrinsic interests in pursuit of intellectual satisfaction (Chamorro-Premuzic & Furnham, 2009; Clark & Schroth, 2010), which mediates the correlation between Openness and academic performance (Swanberg & Martinsen, 2010). One of the advantages of deep learning is the associated pursuit of knowledge even when it is not readily accessible, effectively creating ‘desirable difficulties’ that should enhance learning (Bjork et al., 2013).

1.2 Self- & Other-Rated Personality

Unfortunately, most of the research on personality and educational outcomes has used self-ratings, which have several disadvantages, not least of which being that self-rated personality often has lower correlations with performance than does other-rated personality (Connelly & Ones, 2010). Further, there are times when self-rated personality can be less-than-accurate, such as when there are incentives to manipulate how one is perceived (Kolar, Funder, & Colvin, 1996; Morgeson et al., 2007), or when scales are highly evaluative, meaning they closely reflect overall positive or negative judgments about the target (Kenny & West, 2010).

Vazire (2010) presented a summary of factors that contribute to differences between self- and other-rated personality, including differences with respect to: access to thoughts and feelings; length of observation of the target of personality ratings; and motivational biases when providing personality ratings. Access to relevant observations of the target is one of the major contributors to measurement accuracy (Funder, 1995), and self-raters have unique access to such observations, both because of their privileged access to thoughts and feelings, and the fact that they live with the target of their ratings throughout every moment of their conscious lives. By contrast, other-raters have some access to the thoughts and feelings of their rating target, but this is typically through inference. These points can be taken to imply that self-ratings should be more accurate, but self-raters may in fact be misled by their access to thoughts and feelings, over-emphasizing these when rating personality scales that are behavioral in nature, such as Conscientiousness (Pytlik Zillig et al., 2002). With respect to motivational biases, self-raters will tend to rate themselves more positively on more evaluative traits such as Agreeableness, Conscientiousness and Openness (John & Robins, 1993), which injects spurious variance to personality ratings, thereby lowering correlations (Ziegler & Buehner, 2009). By contrast, other-raters should be less affected by these evaluative biases (Kenny & West, 2010). Other-raters are also particularly attentive to aspects of personality that have pragmatic value to them (Gill & Swann Jr, 2004), a factor that is likely to increase predictive validity (i.e., the ability to statistically predict a criterion variable) when the other-rater values the relevant criterion variable. This is of particular importance for educators, because it raises the possibility that when compared with self-ratings, teacher-rated personality may not only be more obtainable by teachers, it may also be more valid and informative than student self-rated personality measures because of teachers' attention to educational outcomes in their interactions with students.

Taken together, these factors suggest that other-raters, especially teachers, should be better sources for rating at least some personality traits, which given their contribution to educational outcomes could potentially have important consequences for teaching practice and effectiveness. For example, discovery learning approaches (Lee & Anderson, 2013) help students who are higher in Openness to learn, while students lower on Openness are aided by programmed instruction (Orvis, Brusso, Wasserman, & Fisher, 2011), making it valuable for teachers to be able to accurately assess Openness. Further, accurate feedback from teachers to students should provide useful developmental guidance, which in turn could potentially facilitate learning.

However, it is not clear which traits can be validly assessed by teachers. By way of illustration, Conscientiousness is highly evaluative, limiting the accuracy of self-ratings (Vazire, 2010), but many conscientiousness-related behaviors, such as homework and study, occur outside academic settings, beyond the direct observation of teachers (Lubbers et al., 2010) potentially limiting the accuracy of teacher-rated conscientiousness. Similar issues hold with the highly-evaluative, less-observable dimension of Emotional Stability. Without empirical support, it is not possible to identify how best to assess students' personalities in order to provide the best guidance for teachers and students.

Connelly and Ones's (2010) meta-analysis reported estimates of the relationship of academic performance with other-rated measures of the FFM factors, and compared these with previously-obtained meta-analytic correlations with self-rated personality measures. However, Connelly and Ones review used meta-analytic procedures that preclude effective comparison with other recent meta-analyses. Within Poropat's (2009) meta-analysis of academic performance with self-rated FFM measures, all effect sizes came from studies that used measures explicitly designed to measure the FFM dimensions, while Connelly and Ones included correlations with non-FFM measures that were instead logically-assigned to FFM

categories. In doing so, Connelly and Ones used the coding rules of Hough and Ones (2001), but when Hough (1992) used a similar methodology, it resulted in markedly different estimates to those obtained by Poropat. Richardson et al. (2012) used a similar approach to Poropat and the two meta-analyses had similar findings. So, the Connelly and Ones (2010) findings did not provide an effective comparison of correlations of academic performance with self- and other-rated personality.

A further limitation of the Connelly and Ones (2010) report arises from the fact that the relationship between other-rated personality and academic performance was not the primary purpose of their review, which instead provided a much broader analysis of the validity of other-rated personality. This precluded exploration of educationally-relevant moderator effects, such as those due to age and academic level, which Poropat (2009) found to be highly significant. Given the earlier discussions of the value of testing the usefulness of teacher-rated personality, it is also important to consider different types of other-raters as sources of possible moderator effects.

1.3 Conclusion and Predictions

In conclusion, although personality measures based on the FFM were developed without specifically considering educational outcomes, FFM measures have been shown to be reliable statistical predictors of academic performance, and to have meaningful relationships with learning and associated activities. However, the FFM is currently normally assessed using self-ratings, despite the fact that other-ratings used to be more common and much of the work that led to the development of the FFM used other-raters. Given the considerable theoretical and empirical evidence that self- and other-ratings have different relationships with independent outcomes, greater use should be made of other-ratings, especially when self-ratings potentially limit both apparent correlations and practical usefulness of the FFM

within education. It was expected that when compared with self-rated measures, other-rated FFM measures would have stronger correlations with academic performance than those reported by Poropat (2009). Poropat (2009) found significant moderating effects of age and academic level upon these correlations, which were attributed partly to range restriction associated with declining participation in education as students age, and as students with lower levels of achievement cease formal education. Consequently, similar moderating effects were expected with other-rated FFM measures. However, moderating effects were also attributed to changes in the quality of self-assessed personality ratings, a factor that was expected to be irrelevant in this study. The current meta-analysis also allowed an examination of the moderating effect of using different types of other-raters, such as teachers, relatives or peers. Given that teachers observe students mostly within academic contexts and relatives mostly observe students in non-academic contexts, it was expected that teacher-rated personality should be more strongly correlated with academic performance. Testing the relationship between teacher-rated personality and academic performance would provide a basis for developing methods that teachers can use in assessing students, allowing them to both teach individual students more effectively and provide useful guidance to assist student self-understanding and development.

2. Method

Effect sizes used within this meta-analysis were located by searching relevant research databases (i.e., PsycINFO: a database referencing psychology research; ISI Web of Science databases: referencing science, social science, and arts and humanities publications; MEDLINE: lifesciences and biomedical research; ERIC: education publications; and ProQuest Dissertations and Theses: unpublished doctoral and thesis research), using the following search terms and Boolean operators: (academic OR education OR university OR

school) AND (grade OR GPA OR performance OR achievement) AND (personality OR temperament). The reference lists of meta-analyses by Connelly and Ones (2010) and Poropat (2009), as well as the reference lists of articles identified in the database search were also examined to identify any further reported effect sizes.

Studies were only included within this meta-analysis if they provided measures of association between scales that unambiguously measured the FFM and academic performance. One report (Stricker & Rock, 1998) was excluded because it used measures of personality that combined self- and other-ratings, while another (Shiner, 2000) was excluded because the factor scores were based on ratings developed from responses to child and parent interviews. All studies included within the meta-analytic database used separate raters for the personality and academic performance measures, so the results will not have been subject to common method bias, but it should be noted that this may downwardly bias estimates (Kammeyer-Mueller, Steel, & Rubenstein, 2010). In any case, only one study was excluded on the basis that it had reported correlations of ratings of personality and performance provided by the same raters (i.e., Hair, 1999). To enhance comparability with Connelly and Ones (2010), only studies conducted at secondary and tertiary levels of education were included within the database. In two studies (Botwin, Buss, & Shackleford, 1997; Wagerman & Funder, 2007), correlations were reported that related to academic performance at two different academic levels. In these cases, each set of correlations were entered because the measures of academic performance were independent. With few exceptions, the measures of academic performance were GPA, ranging from one semester to an entire study program (degree or high school diploma). The exceptions were Poropat (2005): one course; Ziegler et al. (2010): one assessment. In a previous meta-analysis, Poropat (2009) reviewed the status of these measures of academic performance and concluded that despite their limitations, grades and GPA meet relevant psychometric standards and are suitable for use as criterion variables

in meta-analytic research. Readers who are interested in the issue of the measurement of academic performance are encouraged to read Poropat (2009).

Several studies reported correlations with personality items that were found to reflect the FFM dimensions in independent factor analyses. For the following studies, academic performance-item correlations were combined into weighted averages, with the factor-loadings used as the weighting variable: Correlations with the California Children's Q-Sort (CCQ: Block & Block, 1980) reported by Alvidrez and Weinstein (1999) were averaged based on an independently-conducted exploratory factor analysis (John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994); correlations with academic performance reported by Webb (Adams, Furniss, & DeBow, 1928; Webb, 1915) were averaged using a later exploratory factor analysis (EFA) of his personality measures (Deary, 1996); correlations with items reported by Smith (1969) were averaged using the factor analysis reported by Smith (1967); correlations reported by Tupes (1957) were averaged using the loadings reported by Tupes and Christal (1958).

In most cases, estimates of scale internal reliability for the personality measures, typically Cronbach's (1949) alpha, were obtained from the study reports themselves. If no estimate of reliability was available from the study report, estimates of alpha were obtained from the original validating report for the personality scales or were based on the average meta-analytic alpha for FFM scales reported previously (Viswesvaran & Ones, 2000). Where estimates of measurement reliability for academic performance were available or could be calculated based on reported statistics, study estimates were used. Otherwise, the estimates of alpha for academic performance reported by Bacon and Bean (2006) were used, in line with earlier research (Poropat, 2009). Where studies did not report an average age for participants at the time of academic performance assessment, this was inferred from the year of education using estimates obtained by Poropat (2009). Moderator analyses were conducted

to test for systematic differences between reported effect sizes using similar procedures to those used by Poropat (2009, *in press*), in order to enhance comparability with those reports.

Insert Table 1 around here

Table 1 summarizes the studies that were ultimately included within the meta-analytic database, which included a total of 16 reports and 87 correlations, ranging from 14 to 22 effect sizes and cumulative sample sizes of 2,763 to 5,498, for relationships between academic performance and Openness or Conscientiousness respectively. Even though non-FFM scales were excluded, the meta-analytic sample compares favorably with the corresponding study reported by Connelly and Ones (2010), including nearly three times as many effect sizes and nearly twice as large a total sample size.

3. Results

The obtained correlations for the meta-analytic sample were aggregated using the Hunter and Schmidt random-effects method (Hunter & Schmidt, 2004) but in addition to credibility values, I^2 (Higgins & Thompson, 2002) was used to assess the degree of heterogeneity. Correlations were also converted (Friedman, 1968) to mean differences (d: Cohen, 1988) to provide a more clear, practical implication, as well as to enhance comparability with other meta-analyses. The results of the meta-analysis are presented in Table 2. By Hemphill's (2003) standards, the correlation with Conscientiousness was relatively large, that with Openness was moderate, and the remaining correlations were relatively small. For each FFM dimension, the confidence interval did not include the value for the corresponding estimate for correlations with self-rated personality, confirming that other-rated personality provided more valid statistical prediction of academic performance.

The largest increases in predictive validity were for correlations with Conscientiousness, Emotional Stability and Openness, while the largest increases in variance explained (R^2) were for Conscientiousness $\Delta R^2 = .11$ and Openness $\Delta R^2 = .07$.

Insert Table 2 around here

The significant values for I^2 indicated the existence of non-random variation in the effect sizes, suggesting the presence of systematic moderating effects. Weighted least squares regression was used to test for specific moderators, using the sample size for each study as the weighting variable (Steel & Kammeyer-Mueller, 2002). As explained in the introduction, moderating effects were tested for the mean age of participants, the year of education in which academic performance was assessed (Years 8 to 12 at high school, years 13 to 15 in undergraduate study, year 16 for post-graduate study), the academic level (secondary versus tertiary education), and the source of ratings (i.e., peer, parent, teacher, spouse or referee). In addition, a number of method variables were tested as potential moderators in order confirm whether observed results were due to an artifact: whether FFM ratings were based on a single rater's responses or were the average of responses from multiple raters; the time lag between personality and academic performance rating; the publication year of the original report; or the type of FFM assessment used (based on independently-developed FFM-linked rating scales such as those developed by Goldberg (1992), factor scores for the CCQ, or factor scores from analyses based on data in the originating report). All moderators were tested using the same procedures as those used by Poropat (2009; *in press*). Only one of these moderating variables had significant effects on correlations, with mean age significantly moderating correlations with Conscientiousness $F_{1,21} = 9.47$; $p = .006$; $B = -.57$; $R^2 = .32$. Sample-weighted average correlations within Conscientiousness declined across age cohorts,

with a sharp decline among samples older than 25 (age 14-17, $\rho = .40$; age 18-20, $\rho = .36$; age 21-24, $\rho = .30$; age 25 & over, $\rho = .11$). However, Table 1 shows that all five of the correlations that used samples older than 25 came from just two studies, so this moderating effect may reflect unique methodological features (e.g., neither study used raters who appeared to have observed students in educational settings) and should be interpreted cautiously.

Unexplained heterogeneity along with small-study effects can bias meta-analytic estimates, so Moreno et al. (2012) presented a regression-based method for estimating pooled effect sizes that is robust to these confounds. Moreno et al.'s approach involves regressing effect sizes upon the variance of the effect size, while using the inverse of the effect-size variance as the weighting variable. The intercept with the y-axis (i.e., when the x-value (variance) is theoretically zero) provides the meta-analytic estimate. With the exception of correlations with Openness, this analysis produced estimates that were generally similar to those reported in Table 2 (variance regression-estimated ρ for correlations with: Agreeableness = .11; Conscientiousness = .46; Emotional Stability = .24; Extraversion = .03; Openness = .36). However, the difference between the variance-regressed estimates for the correlations with Conscientiousness and with Openness and those presented in Table 2 (.38 and .28 respectively) indicates that unexplained causes of heterogeneity have suppressed reported correlations with measures of these FFM dimensions.

At the suggestion of anonymous reviewers, meta-analytic correlations of other-rated FFM measures with academic performance were controlled for self-rated FFM and for intelligence. Only one study from this meta-analysis reported cross-correlations for self- and other-rated personality, so instead estimates for self-other correlations for FFM measures were obtained from Connelly and Ones (2010). The sample-weighted average reliability-corrected correlations for family and friend other-raters with self-raters were used, in order to

best reflect the samples in this study (e.g., self-other correlations with incidental acquaintances and strangers were excluded). This produced the following self-other correlations — Agreeableness: .43; Conscientiousness: .54; Emotional Stability: .48; Extraversion: .55; Openness: .53. These values were used to estimate the following part correlations with academic performance once the effect of corresponding self-rated measures had been parceled out: Agreeableness: .08; Conscientiousness: .31; Emotional Stability: .21; Extraversion: .07; Openness: .27. All of the changes in statistical prediction were significant at $p < .001$, indicating significant incremental prediction.

Few studies have reported correlations between intelligence and other-rated measures of the FFM, and only three from within this meta-analysis (Bratko, et al., 2006; Hofmann, 1997; Zimmerman, Triana, & Barrick, 2010) with an aggregate sample of 456, so the following results should be treated cautiously. Openness had the highest sample-weighted, scale-corrected correlations of intelligence of any of the the FFM measures (Agreeableness: . 10; Conscientiousness: .16; Emotional Stability: .12; Extraversion: .05; Openness: .37), which reflects previously argued links between Openness and intelligence. The average of the meta-analytic correlations of intelligence with academic performance reported earlier is .23, and together with the estimates from Table 2 this was used to estimate the following part correlations with academic performance once the effect of intelligence had been parceled out: Agreeableness: .08; Conscientiousness: .35; Emotional Stability: .16; Extraversion: .04; Openness: .21. Apart from Openness, controlling for intelligence had little effect on correlations with academic performance. A hierarchical multiple regression, in which intelligence was entered in the first step and all the FFM measures in the second, produced a substantial increment in statistical prediction $F = 24.14$; $p < .001$; $\Delta R^2 = .19$, equivalent to an increment in d of .97.

4. Discussion

This meta-analysis confirmed that personality variables are important correlates of academic performance, and other-rated personality measures have stronger correlations with academic performance, especially other-rated Conscientiousness, Emotional Stability, and Openness. The strength of these moderating effects is unusually large given “the notorious difficulty of detecting moderator effects” in correlational studies (Shieh, 2009, p. 511).

4.1 Personality and Academic Performance

These findings require a shift in thinking about individual differences and academic performance. The previously observed relative order of correlations of academic performance with self-rated personality (Poropat, 2009) is roughly similar to that reported here for other-rated personality, albeit at a substantially greater magnitude. The overall effect size for other-rated personality was of similar magnitude to some of the largest effect sizes previously reported in education (Hattie, 2009). Consequently, the role of personality requires serious consideration by educators.

Among FFM dimensions, Conscientiousness remains the strongest correlate of academic performance, which is consistent with previous research on the academic habits associated with Conscientiousness. In his ground-breaking research, Webb (1915) proposed a *w* or *will* factor of character to correspond to Spearman’s general intelligence or *g* factor, while Digman (1989) explicitly linked *w* with the third FFM factor, now known as Conscientiousness. In the most comprehensive meta-analysis of correlations between individual differences and academic performance, Richardson et al. (2012) reported only four reliability-corrected correlations that were of greater magnitude to that reported here, and all were specifically linked with academic performance: performance self-efficacy: .67; grade

goal: .49; high school GPA: .41; ACT: .40. So, this research confirms Webb's conjecture with respect to the value of a generic *w* factor within education.

This conclusion is strengthened by comparisons with previous studies. Richardson et al. (2012) reported meta-analytic correlations with *g*-loaded measures ranging from .21 for general intelligence measures to .40 for ACT. These estimates are somewhat lower than those commonly reported for correlations of *g*-loaded measures with tertiary academic performance (.50 to .55: Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008; Neisser et al., 1996; Strenze, 2007), but such estimates are based on range-correction (Kobrin et al., 2008; Neisser et al., 1996) or unrestricted samples (Strenze, 2007), making these estimates non-comparable with either Richardson et al. (2012), Poropat (2009), or the current meta-analysis. This means that the findings are likely to be more similar to those which practitioners will actually observe, at least in those situations where control of range restriction is not possible.

Other-rated Openness is the FFM dimension with the second-strongest correlation with academic performance, apparently because of enhanced curiosity and ability to self-regulate learning. Like Conscientiousness, Openness had stronger correlations with academic performance than did measures of general intelligence in Richardson et al.'s (2012) meta-analysis. The problems with the measurement of this factor outlined in the introduction may account for the fact that correlations with Openness had the largest credibility range. Further research is needed to elucidate the reasons for this degree of variation.

Poropat (2009) reported minimal correlations of self-rated Emotional Stability with academic performance at secondary (.01) and tertiary level (-.01), which were significantly lower than that reported at primary level (.20). To explain this, Poropat (2009) suggested that if it is only the more capable students who continue to higher academic levels, they should be more able to manage negative consequences of low Emotional Stability (e.g., higher levels of anxiety), which would therefore have less effect on academic performance. But the

correlation between Emotional Stability and academic performance reported here (.18) is instead consistent with Poropat's (2009) estimate for this relationship at primary level. So, other-ratings of Emotional Stability appear to maintain their relationship with academic performance at higher educational levels, while self-ratings of Emotional Stability become less-valid as statistical predictors. This conclusion conflicts with previous arguments that other-rated measures of Emotional Stability are less valid than self-rated measures and should be less correlated with relevant criteria (Funder, 2001). Instead, the results of this meta-analysis are encouraging for the use of other-rated measures of this FFM dimension.

This study found a significantly weaker correlation between other-rated Extraversion and academic performance than that reported by Connelly and Ones (2010), but significantly greater than that reported by Poropat (2009). Other-rated Agreeableness was found to have a significant correlation with academic performance, unlike previous meta-analyses, but the effect sizes for both Extraversion and Agreeableness were relatively small.

4.2 Personality Measurement in Education

The strength of the results of this meta-analysis were dependent upon two methodological developments. Firstly, the widespread adoption of lexically-based factors has enabled this type of integrated review of personality in education, and secondly, the recognition that self- and other-ratings are related but not interchangeable makes consideration of rating source more important. However, these results also raise a methodological caution with respect to meta-analyses using the FFM, as demonstrated by comparison with Connelly and Ones (2010). These authors included within their study scales that had not been specifically designed to assess the FFM, and with the exception of Conscientiousness, their correlations were each significantly, and at times substantially, different from those reported in Table 2 (i.e., Agreeableness: .10 & .01; Conscientiousness: .

38 &.41; Emotional Stability: .18 & .27; Extraversion: .05 & .35; Openness: .28 & .18). This contrasts with the previously-noted agreement in estimates between Poropat (2009) and Richardson et al. (2012), who used similar meta-analytic methods. Consequently, meta-analysts need to be careful about how they code scales, but both primary researchers and practitioners should recognize that the results reported here cannot be generalized to non-FFM measures.

It is important to understand the reasons for the differences in correlations with self- and other-rated personality. Vazire (2010) argued that self- to other-raters produce different personality assessments because they have varying access to trait-relevant observations, and contrasting motivations in providing ratings. On this basis, Vazire argued that other-ratings of Emotional Stability would be less accurate reflections of underlying personality, other-ratings of Openness would be more accurate, while self- and other-ratings of Extraversion should be equally accurate, and differential accuracy would be reflected in correlations with criterion variables.

The results presented here do not speak directly to measurement ‘accuracy’, but other-rated measures were consistently more highly-correlated with academic performance, which would appear to be inconsistent with Vazire’s predictions. An alternative explanation is that other-raters attend to information important to the rater (Gill & Swann Jr, 2004), and this information is relevant to academic performance. For other-raters, behaviors are more valuable than thoughts and feelings because behaviors are what others directly experience — thoughts and feelings only affect others through behaviors. But academic performance is a behavioral outcome (Campbell, 1999), and behaviors are best predicted by behaviors (Ouellette & Wood, 1998). Consequently, if Vazire is correct that self-raters have better access to thoughts and feelings, this will reduce the influence on personality ratings by behavioral observations, which would consequently reduce correlations with academic

performance. This raises the possibility that self-ratings can be more valid and accurate reflections of personality as a complex of thoughts, feelings and behaviors, yet have weaker correlations with academic performance.

Consistent with this argument, Lievens, De Corte, and Schollaert (2008) explicitly instructed participants to self-rate personality based upon "how they *behaved* at school" (Lievens, et al., 2008, p. 272: emphasis added to the original), thereby producing a moderating effect on correlations between Conscientiousness and GPA comparable to that reported here (.09 to .37: Lievens, et al., 2008). However, Lievens et al. also asked participants to focus on the school context, and varying the context or frame-of-reference was the primary focus of their research. Moderation by context was in effect tested in this meta-analysis by comparing correlations with personality ratings obtained using different types of other-raters (e.g., parent or teacher), who came from varying contexts. Yet, the type of other-rater was not a significant moderator of correlations. So, some evidential core remains within other-ratings across contexts, making them consistently more highly-correlated with academic performance than are self-ratings. That evidential core appears to be the degree to which other-ratings are based on behavioral evidence, an explanation with readily-applicable implications.

4.3 Implications for Educators

This meta-analysis has shown that when given valid measures, teachers assess students' personalities with some validity, while the research reviewed in the introduction shows that personality is linked with a range of learning-relevant factors. So, teachers should be able to use their own ratings of, for example, students' Conscientiousness in order to improve students' standing on Conscientiousness-related learning factors, such as time-management (Bidjerano & Dai, 2007) and strategic learning (Roberts et al., 2009; Swanberg & Martinsen,

2010). Likewise, if teachers rate students as low on Emotional Stability, they could recommend attentional bias modification training (Dandeneau & Baldwin, 2009) to assist subsequent performance. Accurate assessment of Openness should allow teachers to provide discovery learning to the students who are high on Openness that benefit most from this approach, while presenting more structured learning to students who are low on Openness in order to maximize their learning (Orvis et al., 2011). In other words, other-rated personality should become a valuable guide for educators.

Other-rated personality can also act as a guide to students. As Bjork et al. (2013) argued, students are often ill-informed about what makes learning most effective, which is consistent with the finding that students' self-rated personality is less strongly-associated with academic performance. If students are given personality feedback from teachers and/or peers, this could be used for better understanding and managing their learning-related strengths and limitations. Extending such feedback by carefully integration with behaviorally-based personality self-assessments is likely to both increase self-rating validity, and assist students' ongoing development of learning-management skills (Nowack & Mashihi, 2012).

Bjork et al. (2013) cautioned against "the mindset that one's learning abilities are fixed" (p. 438). Encouragingly, personality has been demonstrated to change over time to a far greater extent than intelligence. The fact that Conscientiousness increases with age (McCrae et al., 2004) impliesthat it may be amenable to development or remediation, offering encouragement to strategies targeted at developing students' related abilities. Promising work has demonstrated the ability to change related factors in childhood, such as endeavours to improve effortful control that increased educational achievement (Barnett et al., 2008; Diamond, Barnett, Thomas, & Munro, 2007). Similar strategies may also improve academic performance by enhancing older students' Conscientiousness, and there is already evidence that Openness (Jackson, Hill, Payne, Roberts, & Stine-Morrow, 2012) and Emotional

Stability (Dandeneau & Baldwin, 2009) can be specifically trained with consequences for subsequent learning. Such efforts have the potential to not only enhance learning, but also to contribute to subsequent employability (Poropat, 2011).

4.4 Limitations

Despite evidence of moderating effects, only one significant moderator was verified, indicating that there is considerably more to be learnt about the relationship between personality and academic performance. A major issue that this study was unable to illuminate is the causal relationship between personality and academic performance. However, the reported correlations do allow researchers and practitioners to use personality measures as indicators of likely outcomes, and to target students accordingly.

4.5 Conclusions

In conclusion, the results of this review confirm that personality is significantly and substantially associated with academic performance, but that personality varies depending on the eye of the beholder, especially if those eyes belong to self- and other-raters. Both types of rater observe something that has a verifiable relationship with independently assessed academic outcomes, but other-ratings are clearly more closely correlated with academic performance. It is intriguing to consider what further research will discover as interactions between raters, targets and criterion variables are explored. One thing is already clear: the role of personality in education is complex, interesting and potent.

References

Adams, H.F., Furniss, L., & DeBow, L.A. (1928). Personality as revealed by mental test scores and by school grades. *Journal of Applied Psychology, 12*(3), 261-277.

Allport, G.W., & Odbert, H.S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs, 47*(1, Whole No. 211).

Alvidrez, J., & Weinstein, R.S. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology, 91*(4), 731-746.

Bacon, D.R., & Bean, B. (2006). GPA in research studies: An invaluable but neglected opportunity. *Journal of Marketing Education, 28*(1), 35-42.

Barnett, W.S., Jung, K., Yarosz, D.J., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*(3), 299-313.

Barrick, M.R., Mount, M.K., & Wonderlic, I. (2006). *Wonderlic Productivity Index.* Libertyville, IL: Wonderlic, Inc.

Bernard, L.C. (2010). Motivation and personality: Relationships between putative motive dimensions and the five factor model of personality. *Psychological Reports, 106*(2), 613-631.

Bidjerano, T., & Dai, D.Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences, 17*(1), 69-81.

Bjork, R.A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417-444.

Block, J., & Block, J.H. (1980). *The California Child Q-set.* Palo Alto, Calif.: Consulting Psychologists Press.

Botwin, M.D., Buss, D.M., & Shackleford, T.K. (1997). Personality and mate preferences: Five factors in mate selection and marital satisfaction. *Journal of Personality, 65*(1), 107-136.

Bowman, M.L. (1989). Testing individual differences in China. *American Psychologist, 44*(3), 576-578.

Bratko, D., Chamorro-Premuzic, T., & Saks, Z. (2006). Personality and school performance: Incremental validity of self- and peer-ratings over intelligence. *Personality and Individual Differences, 41*, 131-142.

Campbell, J.P. (1999). The definition and measurement of performance in the new age. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, motivation, and development* (pp. 399-430). San Francisco: Jossey-Bass.

Cattell, R.B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38*, 476-506.

Chamorro-Premuzic, T., & Furnham, A. (2006). Intellectual competence and the intelligent personality: A third way in differential psychology. *Review of General Psychology, 10*(3), 251-267.

Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly Openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences*, *19*(4), 524-529.

Cianci, A.M., Klein, H.J., & Seijts, G.H. (2010). The Effect of Negative Feedback on Tension and Subsequent Performance: The Main and Interactive Effects of Goal Content and Conscientiousness. *Journal of Applied Psychology, 95*(4), 618-630.

Clark, M.H., & Schroth, C.A. (2010). Examining relationships between academic motivation and personality among college students. *Learning and Individual Differences, 20*, 19-24.

Cohen, J. (1988). *Statistical power analysis of the behavioral sciences.* (2nd ed.). New York: Academic Press.

Connelly, B.S., & Ones, D.S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy & predictive validity. *Psychological Bulletin, 136*(6), 1092-1122.

Costa, P.T., Jr., & McCrae, R.R. (1992). *Revised NEO Personality Inventory and NEO Five-Factor Inventory: Professional Manual*. Odessa Florida: Psychological Assessment Resources Inc.

Cronbach, L.J. (1949). *Essentials of psychological testing.* New York: Harper.

Dandeneau, S.D., & Baldwin, M.W. (2009). The buffering effects of rejection-inhibiting attentional training on social and performance threat among adult students. *Contemporary Educational Psychology, 34*(1), 42-50.

David, A.R. (1998). *Handbook to life in ancient Egypt.* New York: Oxford University Press.

De Pauw, S.S.W., Mervielde, I., & Van Leeuwen, K.G. (2009). How Are Traits Related to Problem Behavior in Preschoolers? Similarities and Contrasts Between Temperament and Personality. *Journal of Abnormal Child Psychology, 37*(3), 309-325.

De Raad, B., & Schouwenburg, H.C. (1996). Personality in learning and education: A review. *European Journal of Personality, 10*, 303-336.

Deary, I.J. (1996). A (latent) Big Five personality model in 1915? A reanalysis of Webb's data. *Journal of Personality and Social Psychology, 71*(5), 992-1005.

DeYoung, C.G., Peterson, J.B., & Higgins, D.M. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality, 73*(4), 825-858.

Diamond, A., Barnett, W.S., Thomas, J., & Munro, S. (2007). The early years - Preschool program improves cognitive control. *Science, 318*(5855), 1387-1388.

Digman, J.M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57*(2), 195-214.

Digman, J.M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*, 116-123.

Duckworth, A.L., Quinn, P.D., & Tsukayama, E. (2012). What no child left behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology, 104*(2), 439-451.

Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4-58.

Edwards, R.E. (1977). Personal traits and "success" in schooling and work. *Educational and Psychological Measurement, 37*(1), 125-138.

Eysenck, H.J., & Cookson, D. (1969). Personality in primary school children: 1. - Ability and achievement. *British Journal of Educational Psychology, 39*, 109-122.

Friedman, H. (1968). Magnitude of experimental effect and a table for its rapid estimation. *Psychological Bulletin, 70*(4), 245-251.

Funder, D.C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.

Funder, D.C. (2001). Accuracy in personality judgement: Research and theory concerning an obvious question. In B. W. Roberts & R. T. Hogan (Eds.), *Personality psychology in the workplace.* (pp. 121-140). Washington DC: American Psychological Association.

Gill, M.J., & Swann Jr, W.B. (2004). On what it means to know someone: A matter of pragmatics. *Journal of Personality and Social Psychology, 86*(3), 405–418.

Goldberg, L.R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26-42.

Graesser, A.C. (2009). Inaugural editorial for *Journal of Educational Psychology*. *Journal of Educational Psychology, 101*(2), 259-261.

Hair, E.C. (1999). *Longitudinal analysis of the self and personality.* (Doctor of Philosophy Doctoral Dissertation), Texas A&M University, College Station, Texas.

Hattie, J.A. (2009). *Visible learning: A synthesis of meta-analyses relating to achievement.* Abingdon, Oxon, UK: Routledge.

Hemphill, J.F. (2003). Interpreting the magnitude of correlation coefficients. *American Psychologist, 58*(1), 78-80.

Hendriks, H.H.J., Hofstee, W.K.B., & De-Raad, B. (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences*, *27*, 307-325.

Higgins, J.P.T., & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558

Hofmann, E.M. (1997). *The relationships between the Five Factor Model of personality and variability in ADHD adolescent functioning.* (Doctor of Philosophy Doctoral Dissertation), University of Akron, Akron Oh.

Hough, L.M. (1992). The "Big Five" personality variables - construct confusion: Description versus prediction. *Human Performance, 5*(1/2), 139-155.

Hough, L.M., & Ones, D.S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil & C. Viswesvaran (Eds.), *Handbook of industrial work and organizational psychology.* (Vol. 1, pp. 233-377). London: Sage.

Hunter, J.E., & Schmidt, F.L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Thousand Oaks, Calif.: Sage.

Jackson, J.J., Hill, P.L., Payne, B.R., Roberts, B.W., & Stine-Morrow, E.A.L. (2012). Can an old dog learn (and want to experience) new tricks? Cognitive training increases openness to experience in older adults. *Psychology and Aging, 27*(2), 286-292.

John, O.P., Caspi, A., Robins, R.W., Moffitt, T.E., & Stouthamer-Loeber, M. (1994). The "Little Five": Exploring the nomological network of the Five-Factor Model of personality in adolescent boys. *Child Development, 65*(1), 160-178.

John, O.P., Donahue, E.M., & Kentle, R. (1991). *The "Big Five" Inventory—Versions 4a and 54*. Berkeley, CA: Institute of Personality Assessment and Research, University of California.

John, O.P., & Robins, R.W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*(4), 521-551.

Kammeyer-Mueller, J., Steel, P.D.G., & Rubenstein, A. (2010). The other side of method bias: The perils of distinct source research designs. *Multivariate Behavioral Research, 45*, 294-321.

Kenny, D.A., & West, T.V. (2010). Similarity and agreement in self- and other-perception: A meta-analysis. *Personality and Social Psychology Review, 14*(2), 196-213.

Kircanski, K., Craske, M.G., & Bjork, R. (2008). Thought suppression enhances memory bias for threat material. *Behaviour Research and Therapy, 46*, 462-476.

Kobrin, J.L., Patterson, B.F., Shaw, E.J., Mattern, K.D., & Barbuti, S.M. (2008). *Validity of the SAT for predicting first-year college grade point average.* New York: The College Board.

Kolar, D.W., Funder, D.C., & Colvin, C.R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality, 64*(2), 311-337.

Kuncel, N.R., & Hezlett, S.A. (2007). Standardized tests predict graduate students' success. *Science, 315*, 1080-1081.

Lee, H.S., & Anderson, J.R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology, 64*, 445-469.

Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology, 93*(2), 268-279.

Lubbers, M.J., Van der Werf, M.P.C., Kuyper, H., & Hendriks, A.A.J. (2010). Does homework behavior mediate the relation between personality and academic performance? *Learning and Individual Differences, 20*(3), 203-208.

MacCann, C., Duckworth, A.L., & Roberts, R.D. (2009). Empirical identification of the major facets of Conscientiousness. . *Learning and Individual Differences, 19*(4), 451-458.

MacCann, C., Fogarty, G.J., & Roberts, R.D. (2012). Strategies for success in education: Time management is more important for part-time than full-time community college students. *Learning and Individual Differences, 22*, 618-623.

MacCann, C., Lipnevich, A.A., Burrus, J., & Roberts, R.D. (2012). The best years of our lives? Coping with stress predicts school grades, life satisfaction, and feelings about high school. *Learning and Individual Differences, 22*(2), 235-241.

McCrae, R.R., Costa, P.T., Hrebickova, M., Urbanek, T., Martin, T.A., Oryol, V.E., . . . Senin, I.G. (2004). Age differences in personality traits across cultures: Self-report and observer perspectives. *European Journal of Personality, 18*(2), 143-157.

Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*(7), 2693-2698.

Moreno, S.G., Sutton, A.J., Thompson, J.R., Ades, A.E., Abrams, K.R., & Cooper, N.J. (2012). A generalized weighting regression-derived meta-analysis estimator robust to small-study effects and heterogeneity. *Statistics in Medicine, 31*, 1407-1417.

Morgeson, F.P., Campion, M.A., Dipboye, R.L., Hollenbeck, J.R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*(3), 683-729.

Mussel, P. (2013). Intellect: A Theoretical Framework for Personality Traits Related to Intellectual Achievements. *Journal of Personality and Social Psychology, 104*(5), 885-906.

Neisser, U., Boodoo, G., Bouchard, T.J.J., Boykin, A.W., Brody, N., Ceci, S.J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77-101.

Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.

Nowack, K.M., & Mashihi, S. (2012). Evidence-based answers to 15 questions about leveraging 360-degree feedback. *Consulting Psychology Journal: Practice and Research, 64*(3), 157-182.

OECD. (2007). *Education at a glance 2007: OECD indicators.* Paris: OECD Publishing.

Orvis, K.A., Brusso, R.C., Wasserman, M.E., & Fisher, S.L. (2011). E-nabled for E-Learning? The moderating role of personality in determining the optimal degree of learner control in an E-Learning environment. *Human Performance, 24*(1), 60-78.

Ouellette, J.A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin, 124*(1), 54-74.

Panadero, E., Tapia, J.A., & Huertas, J.A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences, 22*(6), 806-813.

Poropat, A.E. (2005). *An examination of the relationship between personality and Citizenship Performance in academic and workplace settings.* (Doctoral Dissertation), Griffith University, Brisbane, Qld.

Poropat, A.E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322-338.

Poropat, A.E. (2011). The role of citizenship performance in academic achievement and graduate employability. *Education + Training, 53*(6), 499-514.

Pytlik Zillig, L.M., Hemenover, S.H., & Dienstbier, R.A. (2002). What do we assess when we assess a Big 5 trait?: A content analysis of the affective, behavioral, and cognitive processes represented in Big 5 Personality Inventories. *Personality and Social Psychology Bulletin, 28*(6), 847-858.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 363-387.

Roberts, B.W., Jackson, J.J., Fayard, J.V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior.* (pp. 369-381). New York, NY: Guilford.

Rothbart, M.K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*(4), 207-212.

Rothbart, M.K., & Jones, L.B. (1998). Temperament, self-regulation and education. *School Psychology Review, 27*(4), 479-491.

Saklofske, D.H., Austin, E.J., Mastoras, S.M., Beaton, L., & Osborne, S.E. (2012). Relationships of personality, affect, emotional intelligence and coping with student stress and academic success: Different patterns of association for stress and success. *Learning and Individual Differences*, *22*(2), 251-257.

Saroglou, V., & Fiasse, L. (2003). Birth order, personality, and religion: A study among young adults from a three-sibling family. *Personality and Individual Differences*, *35*(1), 19-29.

Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*, 506-516.

Saucier, G. (2009). Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure. *Journal of Personality, 77*(5), 1577-1614.

Saucier, G., & Goldberg, L.R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical Perspectives.* (pp. 21-50). New York: Guilford.

Saucier, G., & Goldberg, L.R. (2001). Lexical studies of indigenous personality factors: Premises, products and prospects. *Journal of Personality, 69*(6), 847-879.

Saucier, G., Ostendorf, F., & Peabody, D. (2001). The non-evaluative circumplex of personality adjectives. *Journal of Personality, 69*(4), 537-582.

Shieh, G. (2009). Detecting Interaction Effects in Moderated Multiple Regression With Continuous Variables Power and Sample Size Considerations. *Organizational Research Methods, 12*(3), 510-528.

Shiner, R.L. (2000). Linking childhood personality with adaptation: Evidence for continuity and change across time into late adolescence. *Journal of Personality and Social Psychology, 78*(2), 310-325.

Shiner, R.L., & Caspi, A. (2003). Personality differences in childhood and adolescence: measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44*(1), 2-32.

Smith, G.M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement, 27*, 967-984.

Smith, G.M. (1969). *Personality correlates of academic performance in three dissimilar populations.* Paper presented at the 77th Annual Convention of the American Psychological Association.

Spearman, C. (1904). “General intelligence,” objectively determined and measured. . *American Journal of Psychology, 15*, 201-293.

Spearman, C. (1927). *The abilities of man.* London: MacMillan.

Steel, P.D., & Kammeyer-Mueller, J.D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology, 87*(1), 96-111.

Steinmayr, R., Bipp, T., & Spinath, B. (2011). Goal orientations predict academic performance beyond intelligence and personality. *Learning and Individual Differences, 21*(2), 196-200.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*, 401-426.

Stricker, L.J., & Rock, D.A. (1998). Assessing leadership potential with a biographical measure of personality traits. *International Journal of Selection and Assessment, 6*(3), 164-184.

Swanberg, A.B., & Martinsen, O.L. (2010). Personality, approaches to learning and achievement. *Educational Psychology, 30*(1), 75-88.

Tupes, E.C. (1957). Relationships between behavior trait ratings by peers and later officer performance of USAF officer candidate school graduates. *Research Report AFPTRC-TN-57-125*. Lackland Air Force Base, Texas: Air Force Personnel Training & Research Center.

Tupes, E.C., & Christal, R.E. (1958). Stability of personality trait rating factors obtained under diverse conditions. *USAF Wright Air Development Center Technical Note, 16*, 58-61.

Tupes, E.C., & Christal, R.E. (1961). Recurrent personality factors based on trait ratings. *USAF ASD Technical Reports*, 61-97.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge assymetry (SOKA) model. *Journal of Personality & Social Psychology, 98*(2), 281-300.

Viswesvaran, C., & Ones, D.S. (2000). Measurement error in "Big Five factors" personality assessment: Reliability generalization across studies and measures. *Educational and Psychological Measurement, 60*(2), 224-235.

Von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6*(6), 574-588.

Wagerman, S.A., & Funder, D.C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality, 41*, 221-229.

Webb, E. (1915). *Character and Intelligence: An Attempt at an Exact Study of Character.* Cambridge, U.K.: Cambridge University Press.

Wiggins, N. (1969). Prediction of first year success in psychology peer ratings. *Journal of Educational Research, 63*(1), 81-85.

Zhang, X., Nurmi, J.-E., Kiuru, N., Lerkkanen, M.-K., & Aunola, K. (2011). A teacher-report measure of children's task-avoidant behavior: A validation study of the Behavioral Strategy Rating Scale. *Learning and Individual Differences, 21*(6), 690-698.

Zhao, B. (2011). Learning from errors: The role of context, emotion, and personality. *Journal of Organizational Behavior, 32*, 435-463.

Ziegler, M., & Buehner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement, 69*(4), 548-565.

Ziegler, M., Danay, E., Scholmerich, F., & Buhner, M. (2010). Predicting academic success with the Big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality, 24*(4), 341-355.

Ziegler, M., Danay, E., Heene, M., Asendorpf, J., & Buhner, M. (2012). Openness, fluid intelligence, and crystallized intelligence: Toward an integrative model. *Journal of Research in Personality, 46*(2), 173-183.

Zimmerman, R.D., Triana, M.D., & Barrick, M.R. (2010). Predictive criterion-related validity of observer ratings of personality and job-related competencies using multiple raters and multiple performance criteria. *Human Performance, 23*(4), 361-378.

Table 1. Summary of Studies Included in the Meta-Analysis.

Lead Author	Year	FFM Measure	Single-Rater?	FFM Rater	N	Mean Age	Year of Education	Reported FFM Scales
Alvidrez	1999	CCQ	Y	Teacher	96	18	12	ACEmExO
Botwin	1997	Goldberg	Y	Spouse	102	25.5	15	ACEmExO
				Spouse	87	25.5	15	ACEmExO
				Spouse	91	26.7	15	ACEmExO
				Spouse	100	26.7	15	ACEmExO
Bratko	2006	FFPI	Y	Peer	255	16.2	11	ACEmExO
Digman	1989	Digman	Y	Teacher	185	16.9	12	AC
Edwards	1977	Smith	N	Peers	237	16.9	12	ACEmExO
Hofman	1997	CCQ	Y	Parent	94	14.8	10	ACEmExO
Poropat	2005	Saucier	Y	Peer	118	20.9	13	C
Smith	1967	EFA	N	Peers	348	19.2	13	ACEmExO
Smith	1969	Smith	N	Peers	1022	16.9	12	ACEmEx
			N	Peers	798	19.2	13	ACEmEx
Stricker	1998	EFA	N	Peers	348	19.2	13	ACEmEx
Tupes	1957	EFA	N	Peers	790	23.6	16	ACEmExO
Wagerman	2007	BFI	N	Peers	131	16.9	12	C
		BFI	N	Peers	131	19.2	13	C
Webb	1915	EFA	N	Peers	194	19.2	13	ACEmExO
Wiggins	1969	Norman	N	Peers	104	19.2	13	ACEmExO
Ziegler	2010	NEO-PI-R	N	Peers	145	21.9	13	ACEmExO
Zimmerman	2010	Wonderlic	Y	Referee	127	30.0	16	ACEm

Source of Personality Scales: BFI – Big Five Inventory (John, Donahue, & Kentle, 1991); CCQ – California Child Q-Set (Block & Block, 1980), factor-analyzed by John et al. (1994); Digman - Hawaii Scales for Judging Child Behavior (Digman & Inouye, 1986); EFA = Scales based on analysis of data from within the reported study; FFPI = Five Factor Personality Inventory (Hendriks, Hofstee, & De-Raad, 1999); NEO-PI-R – NEO-Personality Inventory-Revised (Costa & McCrae, 1992); Norman = Scales based on Norman (1963); Saucier - Mini-Markers (Saucier, 1994); Smith = Factor-scales provided by Smith (1967); Wonderlic – Wonderlic Productivity Index (Barrick, Mount, & Wonderlic, 2006).

FFM scales: A = Agreeableness; C = Conscientiousness; Em = Emotional Stability; Ex = Extraversion; O = Openness

Table 2. Meta-analysis of self- & other-rated personality measures and academic performance.

FFM dimension							ρ : 95% Confidence Interval		ρ : 80% Credibility Interval		
	k	*N*	*r*	ρ	*d*	ρ_{self}	Lower	Upper	Lower	Upper	I^2
Agreeableness	17	4914	.09	.10	.21	.06	.04	.17	-.15	.35	75.3%***
Conscientiousness	22	5541	.33	.38	.83	.22	.33	.44	.14	.63	84.6%***
Emotional Stability	17	4856	.16	.18	.37	.00	.12	.24	-.04	.41	81.0%***
Extraversion	17	4849	.04	.05	.09	-.02	-.01	.10	-.15	.25	76.2%***
Openness	14	2806	.23	.28	.58	.09	.21	.35	.06	.50	76.2%***

k = number of samples; N = aggregate sample; r = sample-weighted correlation; ρ = sample-weighted correlation corrected for scale reliability; *d* = Cohen's *d*; ρ_{self} = sample-weighted correlation of academic performance with self-rated personality for secondary & tertiary education, corrected for scale reliability (Poropat, 2009); I^2 = index of heterogeneity.

*** $p < .001$