

A meta-analysis of adult-rated child personality and academic performance in primary education.

Abstract

Background:

Personality is reliably associated with academic performance, but personality measurement in primary education can be problematic. Young children find it difficult to accurately self-rate personality, and dominant models of adult personality may be inappropriate for children.

Aims:

This meta-analysis was conducted to determine the validity of the Five-Factor Model (FFM) of personality for statistically predicting children's academic performance.

Sample:

Literature search identified 12 reports, with cumulative sample sizes ranging from 4,382 (19 correlations) to 5,706 (23 correlations) for correlations with Emotional Stability and Conscientiousness respectively.

Method:

Hunter-Schmidt random effects meta-analysis was used and moderators were tested using sample-weighted regression.

Results:

When compared with self-rated measures, adult-rated Conscientiousness and Openness were more-strongly correlated with academic performance, but adult-rated Agreeableness was less-strongly correlated. Q-set based assessments had lower validity, which appeared to explain moderating effects of rating-source. Moderating effects were not found for age, year of education (grades 1 to 7), or language within which the study was conducted.

Conclusions:

Conscientiousness and Openness had two of the strongest correlations with academic performance yet reported, comparable with previous meta-analytic correlations of academic performance with instructional quality, cognitive ability, and feedback. The FFM appears to be valid for educational research with children. Openness, which has no counterpart in models of children's temperament, should be further researched with children. Future research should examine the measurement of childhood personality, its relationship to intelligence, the extent to which it is malleable in primary education, and its causal relationship with academic performance.

Keywords: Other-ratings, Personality, Temperament, Primary Education, Academic Performance, Meta-analysis

Personality has increasingly been recognized as having an important role in education, with recent meta-analyses demonstrating that the dimension of Conscientiousness reliably predicts academic performance in primary education (Poropat, 2009). At this educational level, however, the assessment of personality itself can be problematic, potentially affecting both measurement validity and observed correlations. These problems arise partly from the common practice of assessing personality using self-ratings, and partly from the ongoing development of children's personality, raising questions about the appropriate conceptual models to apply.

Several authors have directly examined children's ability to accurately rate their own personalities. Allik, Laidra, Realo and Pullman (2004) found that when children as old as 12 assessed their own personalities using traditional ratings, the resultant factor structures ratings were less valid, while, Soto, John, Gosling and Potter (2008) observed reduced measurement validity. Such problems seem to account for the substantial moderating effect of age on correlations between academic performance and self-rated personality among primary school children (Poropat, 2009).

Allik et al. (2004) argued that the unreliability of factor structures in children's self-ratings may be due to children's personality structures still being in the process of development. On the basis of similar arguments, Connelly and Ones (2010) restricted their meta-analysis of personality and academic performance to studies with participants older than 14 years of age "to avoid including targets for whom personality development/change was more likely" (p.1098). This caution is inconsistent with the findings of Measelle, John, Ablow, Cowan, and Cowan (2005), who demonstrated acceptable test-retest reliability and criterion-related validity with specially-constructed self-ratings of children as young as five. Likewise, parent-rated personality of children between the ages of 6 and 13 shows respectable levels of stability, cross-informant reliability and factorial invariance (De Fruyt et

al., 2006), so there appears to be something consistent and predictable about children's personalities.

A different problem is presented by the fact that Measelle et al. (2005) and De Fruyt et al. (2006) used assessments based on a personality model that had originally been developed with adults, the Big Five, also known as the Five-Factor Model (FFM). The five dimensions of this model are: Agreeableness (being likable and friendly), Conscientiousness (reliable, dependable, and striving to achieve), Emotional Stability (often referred to by its opposite pole as Neuroticism, and reflecting emotional strength and adjustment versus insecurity and anxiousness), Extraversion (outgoing and sociable), and Openness (also known as Openness to Experience, Intellect and Imagination) (Poropat, 2009).

Despite its reliability (Saucier, 2009), the fact that the FFM was based on a model initially developed with adults raises questions about its appropriateness with children. This concern is comparable with concerns raised by attempts to generalize the FFM from English, the language in which the FFM was developed, to other languages. Recently, De Raad, Barelds, Ostendorf, et al. (2010) found that only three of the FFM dimensions are fully replicable across languages. By extension, the fact that FFM measures have good psychometric qualities in juvenile populations does not establish that the FFM structure is appropriate for assessing children's personality. So, it is important that Goldberg (2001) was able to obtain factor structures that reflected the FFM among adult-ratings of children's personalities.

De Pauw, Mevilde and Van Leeuwen (2009) provided further evidence for the validity of the FFM in childhood by comparing the FFM-based Hierarchical Personality Inventory for Children (HiPIC) scales with measures of temperament. Their analyses revealed important overlaps between the FFM and temperament measures, but the different measures are not redundant (De Pauw et al., 2009). In particular, there seems to be no temperament measure

that corresponds to the Openness dimension in the FFM. Further, De Pauw et al. demonstrated that FFM measures were able to statistically predict problem behaviors in childhood, but the HiPIC Imagination scale, which corresponds to Openness, showed only a minor association, at a level ($r = .11$ in a table reporting over 100 correlations) that suggests it may have arisen due to multiple comparison error (Keselman, Cribbie, & Holland, 1999). Likewise, Measelle et al. (2005) reported that all of the FFM measures had reliable associations with behavioral criteria, except for Openness. This is particularly important because, as Funder (1995) argued, the ability to predict independent criteria is one of the key standards by which dispositional measures are judged, so Openness may not be relevant to children. Given the problems associated with children rating their own personality and the issues with respect to applying the FFM to primary-school age populations, it is important to re-examine the relationship between personality and academic performance in primary education by using measures of children's personality that are provided by adults.

There are good theoretical reasons why the FFM dimensions should be related to academic performance in primary education, but these are as varied as the FFM dimensions themselves. In Poropat's (2009) meta-analysis, self-rated Agreeableness was mostly associated with academic performance at primary level, leading to the suggestion that the relationship of Agreeableness with academic performance was due to the effects of socially-desirable Agreeableness-linked behaviors on teachers' evaluations of students within the more intimate primary school setting. If true, similar or stronger effects should be obtained when adult other-raters assess Agreeableness. Conscientiousness has repeatedly been shown to have the strongest relationships with both work (Barrick, Mount, & Judge, 2001) and academic performance (Poropat, 2009), apparently because of its associations with goal-setting, compliance, concentration, and effort regulation (Poropat, 2009). In children, Conscientiousness is associated with Effortful Control (De Pauw, et al., 2009), and Self-

Control (MacCann, Duckworth, & Roberts, 2009) indicating that Conscientiousness may be the expression of a greater ability to self-regulate, leading to higher levels of time on task and consequently greater learning. Emotional Stability, on the other hand, appears to affect academic performance by the absence of a negative: less emotionally stable students are more readily distracted from their tasks by their emotional state, leading to less learning and poorer performance (De Raad & Schouwenburg, 1996). Eysenck (1992) argued that Extraversion should be related to academic performance because of its association with energy levels. Poropat (2009) found that there was an association between Extraversion and academic performance in primary education, and suggested that this may be due to more outgoing children being more noticeable and thereby getting more attention and positive evaluations. Consistent with this, children who talk more and are more outgoing are rated by teachers as more intelligent (Coplan, Hughes, Bosacki, & Rose-Krasnor, 2011), indicating a degree of halo that may extend to ratings of student performance.

Finally, Openness was described by De Raad and Schouwenburg (1996) as reflecting the 'ideal student', a point echoed in findings that self-rated Openness has been found to be correlated with motivation to learn (Tempelaar, Gijsselaers, van der Loeff, & Nijhuis, 2007), think critically (Bidjerano & Dai, 2007), and adult-ratings of children's Openness is correlated with school competencies of nine- and ten-year-olds (Herzhoff & Tackett, 2012). Yet as mentioned before, Openness is the FFM dimension that has the lowest association with childhood temperament measures (De Pauw et al., 2009), and is also the FFM dimension for which its validity among children has been most seriously questioned (Herzhoff & Tackett, 2012). In Poropat's (2009) meta-analysis, self-rated Openness had the second highest correlations with academic performance at higher levels, but in primary education it was Agreeableness that had that position. Whether this was due to variations in ability to accurately self-assess or to problems with the validity of Openness among children is unclear.

When discussing observed differences in validities at different levels of education, Poropat (2009) suggested that these may be partly due to differences in academic environment. Specifically, primary students typically have one teacher for a year and often have smaller classrooms, giving greater opportunity for personality dimensions such as Agreeableness, Emotional Stability and Extraversion to affect student-teacher relationships, and possibly also influence assessments of academic performance. Social desirability associated with these three dimensions may also be important, resulting in a positive halo effect that may raise teachers' ratings of children's academic performance. If these suggestions are correct, there should be evidence of rater-bias: specifically, parent-rated personality should have lower correlations with academic performance than does teacher-rated personality when teachers provide ratings of both sets of constructs. This is because parents observe children in different settings than do teachers, so their ratings will be based on observations of children that will not have been able to influence teachers' relationships with children, nor teachers' assessments of children's academic performance. Additionally, if teachers have substantial idiosyncratic rating biases these should be reflected in both ratings of children's personality and assessments of children's academic performance.

In summary, self-rated personality has been reliably associated with academic performance in primary education but this association is heavily moderated by age, which is consistent with concerns about the ability of young children to accurately self-assess. FFM personality measures have been developed using a model that emerged from studies of adult personality, and without reference to education generally nor academic performance specifically, so there is a need for further systematic examination of the validity of FFM measures with respect to primary education academic performance. It was expected that in primary education, adult other-rated FFM scales will have stronger correlations with academic performance than do self-rated FFM scales. Further, it was expected that parent-

rated FFM scales would have lower correlations with academic performance than do teacher-rated FFM scales. Consequently, this meta-analysis was conducted to provide a test of both the validity of adult other-rated personality as a statistical predictor of academic performance and of the comparative validity of ratings provided by parents and teachers.

Method

Sample

Relevant studies for use within this meta-analysis were located using the following research databases: PsycINFO; ISI Web of Science; MEDLINE; ERIC; and ProQuest Dissertations and Theses. The following search terms and Boolean operators were used for the search: (academic OR education OR school) AND (grade OR GPA OR performance OR achievement) AND (personality OR temperament). Studies were only included within this meta-analysis if they provided measures of association between scales that unambiguously measured academic performance using teachers' ratings such as grades from within the classroom setting. The one study that was excluded on this basis (Kwok, Hughes, & Luo, 2007) used a standardized, decontextualized measure of academic achievement, so it was decided to exclude this study. All studies included within the database used measures of personality that the authors claimed to be directly assessing the FFM. However, one study was excluded because it used a measure of the FFM dimension Conscientiousness that was difficult to reconcile with conventional Conscientiousness measures (i.e., listening and following directions: Beran, Hughes, & Lupart, 2008). Apart from one study that collected academic performance measures from the previous year (Barbaranelli, Caprara, Rabasca, & Pastorelli, 2003), all of the assessments were administered concurrently.

Asendorpf and van Aken (2003) reported correlations of academic performance with personality rated by both parents and peers, but only correlations with the parent ratings were

used. Mervielde (1992) and Mervielde, Buyst and De Fruyt (1995) appeared to be reports of the same sample, so since Mervielde (1992) reported a more detailed analysis by age, only the estimates from that report were included in this meta-analysis. Roskam, van den Plas-Holper, and de Maere-Gaudissart (2001) reported correlations of academic performance with both parent- and teacher-rated personality, but only the parent-rated personality estimates were used. All studies used students' most recent single-year average performance based on assessments or ratings from their current teacher or school records of assessments provided by the current teacher, except for Barbaranelli et al. (2003), who used academic performance measure obtained from multiple teachers.

The final meta-analytic database included 12 reports, which are listed in Table 1. Cumulative sample sizes ranged from 4,382 (19 effect sizes) for correlations with Emotional Stability, to 5,706 (23 effect sizes) for correlations with Conscientiousness. Only seven effect sizes came from English-speaking samples, with other studies originating in Belgium, France, Germany, Italy and Russia. Necessary data for estimating range restriction were not available, so no attempt was made to correct for this artifact. However, all of the studies came from countries with compulsory primary education, so range restriction was likely to be negligible. Where available, values for Cronbach's (1949) alpha provided in the original report were used to estimate scale reliability. In some cases, authors only reported a range of alpha estimates, so the average of those estimates was used. When reports did not provide any estimate of scale reliability, estimates were obtained from previous research. For personality, estimates from Viswesvaran and Ones's (2000) meta-analysis of personality scale reliability were used, while for academic performance, estimates were based on Bacon and Bean's (2006) analysis of grades and grade point average (GPA) by number of subjects and length of time (e.g., number of terms or years).

Insert Table 1 around here

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

Lead Author	Year	FFM Measure	Language	FFM Rater	N	Mean Age	School Grade	Reported FFM Scales
		CCQ			141	5	1	
Asendorpf	2003	CCQ	German	TCC	124	10	5	ACEmExO
Barbaranelli	2003	GBA BFQ-C	Italian	T	111 520	12 12.4	6 7	CExO
De Fruyt	2008	HiPIC	Flemish	MM	281 318	10.7	6	CO
Hair	1999	Goldberg	English	T	317	11.6	6.5	ACEmExO
John	1994	CCQ	English	P	350	10.2	4	ACEmExO
MacCann	2007	BFQ-C	English	P	340	12.2	7	ACEmExO
				T	280	6	1	
				T	280	7	2	
Mervielde	1992	B5BBS-25	Flemish	T	300	8	3	ACEmExO
				T	300	9	4	
				T	260	10	5	
				T	260	11	6	
					94	7	2	
Roskam	2001	EBMCF	French	M	97	9	4	ACEmExO
					99	9	4	
					109	11	6	
Shiner	2000	PCA	English	P	205	10	4.5	ACEx
Slobodskaya	2007	ICID	Russian	P	606	10.3	6	ACEmExO
Sneed	1989	Digman	English	T	135	9.5	4.5	ACEmExO
Victor	1994	PCA	English	T	179	11.5	5.5	ACEmExO

Personality Scales: BFQ-C = Big 5 Questionnaire for Children (Barbaranelli, et al., 2003); B5BBS-25 = Flemish Big 5 Bipolar Rating Scales (Mervielde, et al., 1995); CCQ = FFM scales based on California Child Q-Set (Block & Block, 1980); Digman = Hawaii Scales for Judging Child Behavior (Digman & Inouye, 1986); EBMCF = Bipolar rating scales based on the FFM (Roskam, Van den Plas-Holper, & De Maere-Gaudissart, 2000); GBA = German Bipolar Adjective Big Five Assessment (Ostendorf, 1990); HiPIC = Hierarchical Personality Inventory for Children; Goldberg = Big 5 Markers

(Goldberg, 1992); ICID = Inventory of Child Individual Differences (Halverson et al., 2003); PCA = Scales based on a principal components analysis.

Personality Rating Source: T = Teacher-rated; M = Mother-rated; P = Parent-rated; C = Care-giver-rated.

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

Results

Hunter and Schmidt's (2004) random effects method was used to estimate the meta-analytic correlations and associated statistics. Credibility intervals and Higgins and Thompson's (2002) I^2 were used to assess heterogeneity. Estimated sample-weighted reliability corrected correlations were also converted to mean differences (d : Cohen, 1988). The results of these analyses are presented in Table 2. When Cohen's (1988) standards for effect sizes are applied, the correlations of academic performance with Conscientiousness can be considered to be relatively large, correlations with Openness are moderate to large, while those with Agreeableness, Emotional Stability and Extraversion were relatively small. According to Hemphill (2003), the correlations with Conscientiousness and Openness lie in the upper third of meta-analytic estimates, while the remaining correlations are within the lower third.

 Insert Table 2 around here

Table 2. Meta-analysis of adult other-rated FFM measures and academic performance.

FFM dimension												
							ρ : 95%		ρ : 90%			
							Confidence		Credibility			
							Interval		Interval			
	k	N	r	ρ	d	ρ_{self}	Lower	Upper	Lower	Upper	I^2	
Agreeableness		4587	.08	.09	.19	.30	.02	.17	-.22	.41	86.4%***	

Conscientiousness	5706	.43	.50	1.14	.28	.43	.56	.19	.80	91.5%***
Emotional Stability	4382	.15	.18	.36	.20	.11	.24	-.08	.44	82.1%***
Extraversion	5107	.10	.11	.23	.18	.06	.16	-.08	.31	72.0%***
Openness	5501	.37	.43	.96	.24	.37	.50	.15	.72	89.4%***

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, corrected for scale reliability (Poropat, 2009); I^2 = index of heterogeneity.

*** $p < .001$

Consistent with expectations, correlations of academic performance with other-rated Conscientiousness and Openness were significantly higher than correlations with corresponding self-rated measures (in both cases, $p < .001$). Contrary to expectations, correlations of academic performance were significantly lower with other-rated than self-rated Extraversion $p = .023$, and Agreeableness $p < .001$.

The substantial credibility ranges and significant values for I^2 indicated the existence of non-random variation and therefore possible moderating effects. Specific effects were examined using weighted least squares regression, with sample sizes used to weight their respective correlations (Steel & Kammeyer-Mueller, 2002). A significant moderating effect was found for the source of personality ratings upon correlations with Conscientiousness $F_{1,21} = 4.33$; $p = .050$; $B = .41$ (correlations with parent-rated Conscientiousness: $\rho = .43$; $d = .95$; $k = 12$; $N = 2,734$; $I^2 = 87.1%$; $p < .001$; correlations with teacher-rated Conscientiousness: $\rho = .56$; $d = 1.34$; $k = 11$; $N = 2,972$; $I^2 = 93.4%$; $p < .001$). The type of scale used in the original study (Q-set, questionnaire, or rating scales) also significantly moderated correlations with Conscientiousness $F_{1,21} = 8.98$; $p = .007$; $B = .55$, and Openness $F_{1,21} = 5.33$; $p = .032$; B

= -.46. For both dimensions, Q-set scales had lower validities (Conscientiousness $\rho = .27$; Openness $\rho = .25$). Three of the four studies that used Q-sets also used parent-ratings, so it was necessary to test whether scale type or rating source explained the moderating effect on correlations with Conscientiousness. A step-wise weighted least squares multiple regression showed that using parents to rate personality had no moderating effect on correlations with Conscientiousness once the moderating effect of using Q-sets to assess personality had been taken into account $\Delta R^2 = .08$; $\Delta F_{1,20} = 2.64$; $p = .120$, but use of Q-sets had a significant moderating effect on correlations with Conscientiousness once the use of parents to rate personality was taken into account $\Delta R^2 = .21$; $\Delta F_{1,20} = 6.79$; $p = .017$. Consequently, the more parsimonious explanation is that Q-sets produced the real moderating effect.

Other factors were also tested as potential moderators but all were found to have non-significant effects on correlations: age; year of education (grades 1 to 7); language within which the study was conducted (i.e., English, Flemish, French, German, Italian or Russian).

Two reports (Shiner, 2000; Sneed, 1989) also cited correlations that allowed a very preliminary assessment of the extent to which the correlations of other-rated personality with academic performance could be accounted for by mental ability, and an anonymous reviewer recommended that analysis of this should be reported. The sample-weighted correlation of intelligence with academic performance based on these two articles ($\rho = .54$; $N = 340$) was consistent with previous estimates of around .50 (Neisser et al., 1996; Spinath, Spinath, Harlaar, & Plomin, 2006). Within the two reports, statistical prediction of academic performance was improved over that provided by intelligence by Conscientiousness part $\rho = .31$; $\Delta R^2 = .09$; $\Delta F_{1,337} = 50.93$; $p = .000$, Emotional Stability part $\rho = .35$; $\Delta R^2 = .12$; $\Delta F_{1,337} = 69.64$; $p = .000$ and Openness part $\rho = .51$; $\Delta R^2 = .26$; $\Delta F_{1,337} = 199.89$; $p = .000$. However, the small size of the aggregate sample suggests caution should be exercised in the interpretation of these results.

Discussion

The results of this meta-analysis demonstrate that Conscientiousness and Openness have substantial value as statistical predictors of academic performance in primary education. The magnitude of the correlations with adult other-rated Conscientiousness (corrected $r = .50$; $d = 1.14$) and Openness (corrected $r = .43$; $d = .96$) are much greater than the average of student-linked factors in education ($d = .40$: Hattie, 2009). By way of comparison, intelligence has come to be accepted by many as ‘the dominant predictor of school achievement’ (Spinath, et al., 2006: p.364) on the basis of correlations of around .50 with academic performance (Neisser, et al., 1996; Spinath, et al., 2006), while socio-economic status is only correlated with academic performance at .32 (Sirin, 2005). In several reports, Hattie (1999, 2009; Hattie & Timperley, 2007) has summarized research on learning outcomes reviewed in previous meta-analyses. Some of the largest effect sizes were for provision of learning cues ($d = 1.10$: Hattie & Timperley, 2007), students’ prior cognitive ability ($d = 1.04$: Hattie, 1999), instructional quality ($d = 1.00$: Hattie, 1999), feedback ($d = .95$: Hattie & Timperley, 2007), formative evaluation ($d = .69$: Hattie & Timperley, 2007) and meta-cognitive skills ($d = .69$). This summary indicates that Conscientiousness and Openness should be accepted as important factors within primary education.

Conscientiousness remains the FFM dimension with the strongest association with academic performance, apparently reflecting its basis in effortful control and self-regulation, and its association with motivation, focused effort, and school competence, as reviewed in the introduction. Openness was almost as strongly linked as Conscientiousness to academic performance meaning that even among children, Openness measures are valid. With respect to behavioral problems, De Pauw et al. (2009) argued that measures of childhood temperament and childhood FFM measures may overlap but each add value to the other. The temperament models that De Pauw et al. examined did not have an equivalent to Openness,

so the correlations reported here indicate that temperament theorists may have overlooked an important factor in children's functioning.

The finding that Conscientiousness and Openness are the most important personality factors in primary education and have similar correlations with academic achievement is substantially different from the findings of Connelly and Ones (2010), for whom Openness was much less strongly correlated with academic performance than was Conscientiousness (.18 versus .41) and in fact had the second lowest correlation of the FFM dimensions. This may be a consequence of the different population investigated by Connelly and Ones, which was restricted to post-primary education.

Although the correlations with Conscientiousness and Openness were consistent with the idea that other-rated personality measures will be more valid, the correlations with the other FFM dimensions were not. Adult-rated Extraversion and Agreeableness had lower correlations with academic performance than that observed with self-ratings, and adult-rated Emotional Stability did not have a higher correlation with academic performance than did self-ratings. Blackman and Funder (2002) argued with respect to adults that other-rated measures of Emotional Stability are likely to be less valid because of lower access to the ratee's emotional state. This should in turn result in lower correlations with other variables, which was not observed, so adult-ratings of children's Emotional Stability appear to be as valid as self-ratings. This indicates that adult-raters have good access to emotional states that are relevant to children's academic performance, apparently because children in primary-education have yet to learn how to consistently restrict the ability of others to observe their emotions by learning how to control the expression of their affect. Until they do so, children's Emotional Stability is likely to remain a 'good trait' (i.e., readily observable) in Blackman and Funder's (2002, p.113) terms.

The findings of this research are inconsistent with Poropat's (2009) suggestion that higher levels on Extraversion, Agreeableness and Emotional Stability would have a bigger impact on academic performance in primary education because of either their positive consequences for the relationships between children and their teachers, or because of a social-desirability linked halo effect on measures of academic performance. Despite teachers associating extraverted behavior with intelligence (Coplan, et al., 2011), the modest correlation of teacher-rated Extraversion with academic performance and the lack of a teacher-linked moderating effect show that this association does not have a strong impact on the grades teachers assign to students. More generally, Roskam et al. (2001) specifically examined the relationship between social desirability and their other-rated personality measures and found them to be uncorrelated.

This does not rule out other types of halo effect but they are unlikely to be along the lines that Poropat (2009) suggested. For example, it may instead be that children who perform better at school see themselves as more extraverted, agreeable and emotionally stable, possibly as a consequence of experiencing school as a more rewarding place. An alternative explanation is that children's self-ratings are affected by their academic ability, with children who are performing better at school being more able to accurately comprehend and respond to personality assessments, and in turn being quicker to recognize how best to describe themselves for self-presentation purposes. Such an explanation would be consistent with the substantial moderating effect of age on correlations of academic performance with self-rated Extraversion and Emotional Stability that Poropat (2009) reported.

Evidence for a different type of effect also failed to emerge: whether personality ratings were provided by parents and teachers had no significant differences on correlations associated with varying these sources, despite the differential access and contribution to ratings of children's academic performance. The average corrected inter-correlation of other-

ratings from within a person's own family is only .43 for Conscientiousness and .47 for Openness (Connelly & Ones, 2010), and given the substantial differences between home and school environments it is likely that parent-teacher agreement on personality ratings would be lower. Despite this, both parent- and teacher-rated Conscientiousness and Openness are correlated at similar levels with primary school academic performance. It would be valuable, therefore, for future researchers to conduct multi-rater, multivariate analyses of these relationships that explore what factors parents and teachers are both observing that contribute to academic performance.

Future researchers should also examine the factors that contribute to the significant amount of unexplained between-study variation reported in Table 2. It is inherently difficult to isolate moderator effects in observational studies (Shieh, 2009), so some forms of experimental manipulation may be helpful. On the other hand, the difficulty of isolating moderators makes the substantial effect resulting from using other-raters rather than self-raters that much more important, something that future researchers on children's education and personalities must take into account. For example, it would appear that researchers should seriously question the value of self-rated personality measures with children and only use them when a valid method of obtaining children's self-ratings is available, or if there is a compelling reason to do so such as the desire to explore children's self-perceptions. An example of this is provided by Laidra, Pullman, and Allik (2007), who were specifically interested in children's ability to accurately self-assess. Otherwise, personality researchers and practitioners should take their lead from temperament researchers and rely on other-rated measures when researching children.

The strength of the correlations of academic performance with Conscientiousness and Openness highlights questions relating to the causal relationships between these FFM dimensions and academic performance. Longitudinal studies may assist with teasing out

these relationships, but researchers should also explore the association of both of these personality factors with intelligence. In their article, Asendorpf and van Aken (2003) reported correlations of IQ with adult-rated Conscientiousness that ranged from .03 to .46, and with adult-rated Openness (called 'Culture' in their study) ranging from .32 to .52. This raises the possibility of substantial confounding, especially given that intelligence has long been closely associated with academic performance (Chamorro-Premuzic & Furnham, 2006). The analysis of results reported by Shiner (2000) and Sneed (1989) suggest the existence of some degree of overlap but there remained significant part correlations between academic performance and Conscientiousness, Emotional Stability and Openness after controlling for intelligence. These results should be considered to be preliminary and further studies are needed to test these relationships.

A different but related set of questions relates to the measurement of personality among children. When an adult assesses a child's personality, they make many choices about what observations to attend to and how to combine those into coherent responses. The relationships between the child's actions, the adult's observations, and the ultimate ratings are complex and ill-described, and need to be investigated in order to properly account for the reported correlations. This is more than just a measurement issue, because explanations of these ratings have implications for causal relationships between personality and academic performance. For example, it could be that: adults rate the personality of children differently based on their academic performance; children perform differently based on how adults talk about (and subsequently rate) them; or the behaviors of children that contribute to better academic performance also result in different adult ratings of their personalities. Finally, there is evidence that personality changes and develops with age (McCrae et al., 2004), and factors associated with Conscientiousness, such as self-control and self-discipline, are amenable to development (Heckman & Masterov, 2007). It will be important, therefore, to explore the

degree to which these personality factors can be deliberately developed in order to assist children with their education and later development and success, such as by overtly training students in self-control and self-discipline, or using subtler techniques such as cognitive bias modification to affect Emotional Stability (Macleod & Mathews, 2012). If this is possible, it should provide an opportunity for an especially strong test of the causal relationship between personality and academic performance by using experimental manipulations.

It is more than ten years since Heckman and Rubinstein (2001) expressed surprise at the emphasis in discussions of skill formation upon cognitive ability at the expense of non-cognitive factors. Much of the reason for this has been the lack of reliable evidence. The advent of coherent models for measuring non-cognitive factors has enabled meta-analyses to provide a broader empirical basis (Poropat, 2009), but the emphasis on self-ratings of personality has limited the observed validities (Connelly & Ones, 2010). It will be intriguing to watch over the next few years as researchers endeavor to clarify the role of personality within primary education while using other-ratings, especially by considering its relationship to temperament and exploring methods for developing and applying personality concepts within elementary school environments.

(Asendorpf & van Aken, 2003; Barbaranelli, et al., 2003; De Fruyt, van Leeuwen, de Bolle, & de Clercq, 2008; Hair, 1999; John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994; MacCann, 2007; Mervielde, 1992; Roskam, et al., 2001; Shiner, 2000; Slobodskaya, 2007; Sneed, 1989; Victor, 1994).

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Allik, J., Laidra, K., Realo, A., & Pullman, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality, 18*, 445-462. doi: 10.1002/per.524
- Asendorpf, J. B., & van Aken, M. A. G. (2003). Validity of Big Five personality judgments in childhood: A 9 year longitudinal study. *European Journal of Personality, 17*(1), 1-17. doi: 10.1002/per.460
- Bacon, D. R., & Bean, B. (2006). GPA in research studies: An invaluable but neglected opportunity. *Journal of Marketing Education, 28*(1), 35-42. doi: 10.1177/0273475305284638
- Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences, 34*, 645-664. doi: 10.1016/S0191-8869(02)00051-X
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next. *International Journal of Selection and Assessment, 9*(1-2), 9-30. doi: 10.1111/1468-2389.00160
- Beran, T. N., Hughes, G., & Lupart, J. (2008). A model of achievement and bullying: Analyses of the Canadian National Longitudinal Survey of Children and Youth data. *Educational Research, 50*(1), 25-39. doi: 10.1080/00131880801920379
- 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. doi: 10.1016/j.lindif.2007.02.001
- Blackman, M. C., & Funder, D. C. (2002). Effective interview practices for accurately assessing counterproductive traits. *International Journal of Selection and Assessment, 10*(1/2), 109-116. doi: 10.1111/1468-2389.00197
- Block, J., & Block, J. H. (1980). *The California Child Q-set*. Palo Alto, Calif.: Consulting Psychologists Press.
- 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. doi: 10.1037/1089-2680.10.3.251
- 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*(3), 1092-1122. doi: 10.1037/a0021212
- Coplan, R. J., Hughes, K., Bosacki, S., & Rose-Krasnor, L. (2011). Is silence golden? Elementary school teachers' strategies and beliefs regarding hypothetical shy/quiet and exuberant/talkative children. *Journal of Educational Psychology, 103*(4), 939-951. doi: 10.1037/a0024551
- Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper.
- De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality & Social Psychology, 81*(3), 538-552. doi: 10.1037/0022-3514.91.3.538

- De Fruyt, F., van Leeuwen, K., de Bolle, M., & de Clercq, B. (2008). Sex differences in school performance as a function of conscientiousness, imagination and the mediating role of problem behaviour. *European Journal of Personality, 22*(2), 167-184. doi: 10.1002/per.675
- 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. doi: 10.1007/s10802-008-9290-0
- De Raad, B., Barelds, D. P. H., Ostendorf, F., Mlačić, B., Di Blas, L., Hřebíčková, M., . . . Perugini, M. (2010). Only three factors of personality description are fully replicable across languages: A comparison of 14 trait taxonomies. *Journal of Personality & Social Psychology, 98*(1), 160-173. doi: 10.1037/a0017184
- De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality, 10*, 303-336. doi: 10.1002/(SICI)1099-0984(199612)10:5<303::AID-PER262>3.0.CO;2-2
- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*, 116-123. doi: 10.1037/0022-3514.50.1.116
- Eysenck, H. J. (1992). Personality and education: The influence of Extraversion, Neuroticism and Psychoticism. *German Journal of Educational Psychology, 6*, 133-144.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670. doi: 10.1037/0033-295X.102.4.652
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26-42. doi: 10.1037/1040-3590.4.1.26
- Goldberg, L. R. (2001). Analyses of Digman's child-personality data: Derivation of Big Five factor scores from each of six samples. *Journal of Personality, 69*(5), 709-743. doi: 10.1111/1467-6494.695161
- Hair, E. C. (1999). *Longitudinal analysis of the self and personality*. Doctor of Philosophy Doctoral Dissertation, Texas A&M University, College Station, Texas.
- Halverson, C. F., Havill, V. L., Deal, J. E., Baker, S. R., Victor, J. B., Pavlopoulos, V., . . . Wen, L. (2003). Personality structure as derived from parental ratings of free descriptions of children: the Inventory of Child Individual Differences. *Journal of Personality, 71*(6), 995-1026. doi: 10.1111/1467-6494.7106005
- Hattie, J. A. (1999). Influences on student learning (Inaugural professorial address, University of Auckland, New Zealand). Retrieved from <http://www.education.auckland.ac.nz/webdav/site/education/shared/hattie/docs/influences-on-student-learning.pdf>
- Hattie, J. A. (2009). *Visible learning: A synthesis of meta-analyses relating to achievement*. Abingdon, Oxon, UK: Routledge.
- Hattie, J. A., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi: 10.3102/003465430298487
- Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Review of Agricultural Economics, 29*(3), 446-493. doi: 10.1111/j.1467-9353.2007.00359.x
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review, 91*(2), 145-149. doi: 10.1257/aer.91.2.145
- Hemphill, J. F. (2003). Interpreting the magnitude of correlation coefficients. *American Psychologist, 58*(1), 78-80. doi: 10.1037/0003-066X.58.1.78

- Herzhoff, K., & Tackett, J. L. (2012). Establishing construct validity for Openness-to-Experience in middle childhood: Contributions from personality and temperament. *Journal of Research in Personality, 46*, 286-294. doi: <http://dx.doi.org/10.1016/j.jrp.2012.02.007>
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539-1558 doi: 10.1002/sim.1186
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, Calif.: Sage.
- 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. doi: 10.1111/j.1467-8624.1994.tb00742.x
- Keselman, H. J., Cribbie, R., & Holland, B. (1999). The pairwise multiple comparison multiplicity problem: An alternative approach to familywise and comparisonwise Type I error control. *Psychological Methods, 4*, 58-69. doi: 10.1037/1082-989X.4.1.58
- Kwok, O. M., Hughes, J. N., & Luo, W. (2007). Role of resilient personality on lower achieving first grade students' current and future achievement. *Journal of School Psychology, 45*(1), 61-82. doi: 10.1016/j.jsp.2006.07.002
- Laidra, K., Pullman, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences, 42*, 441-451. doi: 10.1016/j.paid.2006.08.001
- MacCann, C. (2007, July 22nd - July 27th). *Noncognitive constructs in middle schools*. Paper presented at the XIIIth Biennial Meeting of the International Society for the Study of Individual Differences, Giessen, Germany.
- MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19*, 451-458. doi: 10.1016/j.lindif.2009.03.007
- Macleod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Psychology, 8*, 189-217. doi: 10.1146/annurev-clinpsy-032511-143052
- McCrae, R. R., Costa, P. T., Hrebickova, M., Urbaneck, T., Martin, T. A., Oryol, V. E., . . . Senin, I. G. (2004). Age differences in personality traits across cultures: Self-report and observer perspectives. [Proceedings Paper]. *European Journal of Personality, 18*(2), 143-157. doi: 10.1002/per.510
- Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality & Social Psychology, 89*(1), 90-106. doi: 10.1037/0022-3514.89.1.90
- Mervielde, I. (1992). The B5BBS-25: A Flemish set of bipolar markers for the Big-5 personality factors. *Psychologica Belgica, 32*(2), 195-210.
- Mervielde, I., Buyst, V., & De Fruyt, F. (1995). The validity of the Big-Five as a model for teachers' ratings of individual differences among children aged 4-12 years. *Personality and Individual Differences, 18*(4), 525-534. doi: 1854/LU-200615
- 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. doi: 10.1037/0003-066X.51.2.77
- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit. [Language and personality structure: Validity of the five-factor model of personality.]*. Regensburg: Roderer.

- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322-338. doi: 10.1037/a0014996
- Roskam, I., Van den Plas-Holper, C., & De Maere-Gaudissart, A. (2000). Mise au point d'un instrument d'évaluation de la personnalité des enfants à partir du Modèle à Cinquante facteurs [Assessment of the child's personality: Validation of bipolar rating scales based on the Five Factor Model]. . . *L'Orientation Scolaire et Professionnelle*, 29(4), 661-672.
- Roskam, I., van den Plas-Holper, C., & de Maere-Gaudissart, A. (2001). Mothers' and teachers' ratings of children's personality: Children's age, gender, scholastic performance, mothers' educational level, and rater effects. *European Review of Applied Psychology*, 51(4), 289-303.
- Saucier, G. (2009). Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure. *Journal of Personality*, 77(5), 1577-1614. doi: 10.1111/j.1467-6494.2009.00593.x
- 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. doi: 10.1177/1094428108320370
- 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. doi: 10.1037//0022-3514.78.2.310
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417-453. doi: 10.3102/00346543075003417
- Slobodskaya, H. R. (2007). The associations among the Big Five, Behavioural Inhibition and Behavioural Approach systems and child and adolescent adjustment in Russia. *Personality and Individual Differences*, 43(4), 913-924. doi: 10.1016/j.paid.2007.02.012
- Sneed, T. G. (1989). *The relationship of selected personality variables to reading, math, and language arts achievement in 4th and 5th grade students*. Doctor of Philosophy Doctoral Dissertation, University of California, Riverside.
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from Ages 10 to 20. *Journal of Personality and Social Psychology*, 94(4), 718-737. doi: 10.1037/0022-3514.94.4.718
- Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, 34, 363-374. doi: 10.1016/j.intell.2005.11.004
- 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. doi: 10.1037//0021-9010.87.1.96
- Tempelaar, D. T., Gijsselaers, W. H., van der Loeff, S. S., & Nijhuis, J. F. H. (2007). A structural equation model analyzing the relationship of student achievement motivations and personality factors in a range of academic subject-matter areas. *Contemporary Educational Psychology*, 32(1), 105-131. doi: 10.1016/j.cedpsych.2006.10.004
- Victor, J. B. (1994). The Five-Factor model applied to individual differences in school behavior. In C. F. J. Halverson, G. A. Kohnstamm & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood*. (pp. 355-366). Hillsdale, NJ: Lawrence Erlbaum Associates.

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. doi: 10.1177/00131640021970475