Coversheet

This is the accepted manuscript (post-print version) of the article. Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication
Please cite the final published version:


Publication metadata

Title: The Big Five and tertiary academic performance: A systematic review and meta-analysis
Author(s): Vedel, A.
Journal: Personality and Individual Differences
DOI/Link: 10.1016/j.paid.2014.07.011
Document version: Accepted manuscript (post-print)
Anna Vedel
Department of Psychology and Behavioural Sciences,
Aarhus University, Denmark

The Big Five and tertiary academic performance: A systematic review and meta-analysis
Abstract

This article reports a systematic review and meta-analysis of the relationships between the Big Five personality traits and tertiary academic performance. Five frequently used personality measures formed a restricted inclusion criteria pertaining to predictor variables: NEO Personality Inventory – Revised (NEO-PI-R; Costa & McCrae, 1992), NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992), Big Five Inventory (BFI; John, Naumann, & Soto, 2008), Big Five Markers (Goldberg, 1992; Saucier, 1994; Thompson, 2008), and Big Five International Personality Item Pool (IPIP, www.ipip.ori.org/). Grade point average (GPA) was the criterion variable. 20 studies (21 independent samples) published between 1996 and 2013 were included with a total of 105 correlations and an aggregated sample size of 17,717. A random-effects model was used for meta-analysis. GPA was found to correlate significantly with Agreeableness, Conscientiousness, and Openness. Conscientiousness was the strongest predictor of GPA by far with a weighted summary effect of .26. Subgroup analyses tested a potential moderator variable not explored hitherto: academic major of study participants. Academic major was indeed found to moderate the relationship between Conscientiousness and GPA. Problems with the widespread use of psychology students only in samples and other methodological issues are discussed, and suggestions are provided for future research.

Keywords: personality; Big Five; academic performance; GPA; meta-analysis; moderation
1. Introduction

In the philosophical writings of Plato the concept *reason* is what distinguishes the academically capable from others (Taylor, Reeves, & Jeffords, 2008), and historically the search for predictors of superior academic performance has primarily focused on cognitive abilities (Sternberg, 1990). A milestone in this search was the workings of Alfred Binet from the beginning of the 20th century. Binet studied school children’s differential academic performance and constructed intelligence tests and scales in order to identify pupils with academic potential (Binet & Simon, 1916). In the same historical time period Charles Spearman published work pointing towards a general intelligence factor (Spearman, 1904), a factor that has since become widely known as the $g$ factor. Binet and Spearman’s workings had an enormous impact, and cognitive abilities have consistently been shown to predict academic performance well (Ackerman & Heggestad, 1997). But there always were researchers doubting the adequacy of solely focusing on cognitive abilities. Edward Webb studied individual differences and academic performance of British college students and schoolboys and concluded that “character” was just as important as intelligence in predicting academic success (Webb, 1915). Webb even proposed a general will-factor, $w$, corresponding to Spearman’s $g$ to emphasize the role of personality in academic performance. However, the $w$ factor never gained currency, and cognitive factors remained the focus of research in prediction of academic performance throughout the century. With the growing consensus on the Big Five personality traits, *Agreeableness, Conscientiousness, Extraversion, Neuroticism,* and *Openness* (McCrae & Costa, 2008), however, research into the personality-academic performance interface has become more frequent.
1.1. Variability in Big Five measures in previous research

Three meta-analyses have reviewed the broad literature on the predictive validity of the Big Five for academic performance (Poropat, 2009; Richardson, Abraham, & Bond, 2012; Trapmann, Hell, Hirn, & Schuler, 2007) and found substantial, positive correlations between Conscientiousness and academic performance (grades and GPA) in the range of .19 - .27. When partialing out intelligence whenever possible, the correlation between Conscientiousness and academic performance increases (e.g. Poropat, 2009); establishing Conscientiousness as an independent predictor not confounded with psychometric intelligence. Only small correlations were found for the other four personality traits. These meta-analyses had very broad inclusion criteria; studies using any measure of the Big Five personality traits were included. However, it is plausible that the various measures of the Big Five are not equally good at predicting academic success. Though statistical corrections were applied regarding sampling error (Richardson et al., 2012) and reliability (Poropat, 2009) whenever possible, the results still are based on collapsing diverse measures leading to a somewhat blurred picture.

1.2. Variability in criterion in previous research

This potential problem in criterion-related validity was recently addressed by McAbee and Oswald (2013) who performed a meta-analysis on the criterion-related validity of five frequently used personality measures for predicting tertiary level academic performance (grades and GPA): the NEO Personality Inventory – Revised (NEO-PI-R, Costa & McCrae, 1992), the NEO Five-Factor Inventory (NEO-FFI, Costa & McCrae, 1992), the Big Five Inventory (BFI, John, Naumann, & Soto, 2008), Big Five Markers (Goldberg, 1992; Saucier, 1994; Thompson, 2008), and the Big
**Five International Personality Item Pool (IPIP, www.ipip.ori.org).**

Conscientiousness was found to be a robust predictor of academic performance across these particular five personality measures adding to the validity of future meta-analyses collapsing any of these particular five personality measures. But the criterion variable in the meta-analysis by McAbee and Oswald was very broad and covered both registry GPA, self-reported GPA, and individual course grades. Studies have generally found very high correlations between self-reported GPA and registry GPA in the range .84 - .89 (Gray & Watson, 2002; Noftle & Robins, 2007), which makes self-reported GPA a suitable criterion variable comparable to registry GPA. Individual course grades on the other hand is much more problematic. Courses vary greatly in content and kinds of academic assessment (Furnham & Chamorro-Premuzic, 2005), and course grades are not comparable with GPA due to their specificity.

### 1.3. Aims of the current meta-analysis

The first objective of the current meta-analysis was to find out to which extend more restricted selection criteria of Big Five personality measures as well as criterion variable yield higher predictive validity of the Big Five, Conscientiousness especially, for academic performance in college and university populations, compared to the four meta-analyses previously undertaken with more variability in personality measures and criterion variable.

The second objective was to assess a new, potential moderator of these relationships: the academic major of the subjects participating in the studies. Previous research (Poropat, 2009) has assessed the impact of moderators such as age and educational level but never explored whether there are differences in the strength of
association between the Big Five personality traits and academic performance depending on the academic major of study participants. Most studies in this field of research use samples consisting of psychology students only. It is relevant to investigate whether these studies yield results similar to or different from studies performed with students from other academic fields since discrepancies would have implications for the generalizability of previous findings on the predictive validity of the Big Five personality traits, Conscientiousness especially, for academic performance. The current meta-analysis is the first to test this potential moderator.

2. Method

2.1. Literature search

A systematic search by thematically relevant electronic databases was conducted to identify studies on the relationship between the Big Five personality traits and academic performance at university. Using ProQuest the following electronic databases were searched simultaneously with the last search run on April 14, 2014: Australian Education Index (1977 – present), British Education Index (1975 – present), ERIC (1966 – present), PsycINFO (1806 – present) and Sociological Abstracts (1952 – present). Search terms and Boolean operators were ab(personality) AND ab(academic success OR academic performance) AND peer(yes). No publication date limits were applied. Abstracts of the located studies were reviewed, and potential relevant studies were identified. Copies of the potential relevant studies were obtained and examined applying the inclusion criteria outlined in the section below. References of included studies were searched manually to identify additional relevant studies. Finally, this author’s personal collection of electronic articles in the
personality research field was searched identifying a relevant, recent study (Furnham, 2012) not cited in the included studies or in any of the previous meta-analyses.

2.2. Inclusion criteria

The current meta-analysis included only studies available in English for further examination. The most basic requirement for inclusion of studies was that one of the personality measures shown to have equally good criterion-related validity (McAbee & Oswald, 2013) had been correlated with GPA\textsuperscript{1}. Studies using other personality measures than the NEO-PI-R, NEO-FFI, BFI, Markers, or IPIP were therefore excluded (e.g. Busato, Prins, Elshout, & Hamaker, 2000; Cela-Ranilla, Gisbert, & de Oliveira, 2011; Paunonen & Ashton, 2013; Ridgell & Lounsbury, 2004). In one study (Gray & Watson, 2002) Conscientiousness, was measured by the NEO-PI-R whereas the remaining four factors were measured by the NEO-FFI. Given that both of these personality measures honour the inclusion criteria this study was included. Studies employing only a subset of one of the personality measures were not included (e.g. Ahmad, 2011; Furnham, Nuygards, & Chamorro-Premuzic, 2013; Wagerman & Funder, 2007), and studies using single course or examination grade as criterion variable were also excluded (e.g. Ziegler, Danay, Schölmerich, & Bühner, 2010). In the study conducted by Ferguson, Sanders, O’Hehir, and James (2000) the criterion variable was a so-called “general medical training factor” extracted with exploratory factor analysis from 21 assessment variables and yielding a coefficient alpha of .87. The generality yet cohesion - as shown by factor analysis - of this factor

\textsuperscript{1} In Europe the academic performance criterion typically used is overall exam marks not encompassing coursework as the American GPA. For the sake of simplicity GPA will be used in the current meta-analysis to denounce both.
makes it comparable with GPA; therefore this study was included. Studies using samples drawn from the non-tertiary level of education were excluded (e.g. Bratko, Chamorro-Premuzic, & Saks, 2006; Di Fabio & Busoni, 2007).

The identified relevant studies were included in the meta-analysis only if they contained zero-order correlations (or data that could be converted to this) between the Big Five personality traits and GPA. The reason for this inclusion criterion is the diversity in statistical analyses found in the literature. Using zero-order correlations as a frame of reference enables comparison of studies using different statistical strategies such as multiple regression and structural equation modelling. Two studies (Blickle, 1996; Lievens, Coetsier, De Fruyt, & De Maeseneer, 2002) met the inclusion criteria except that they did not provide zero-order correlations. The corresponding authors of these studies were contacted in order to obtain the zero-order correlations, but retrieving these was not deemed feasible by the corresponding authors (G. Blickle, personal communication, November 29, 2013; F. Lievens, personal communication, December 1, 2013). These two studies were therefore excluded. One study reported correlations by gender only (Richardson & Abraham, 2009). In this case total correlations were calculated using the Fisher’s $z$ transformation for averaging correlation coefficients advocated by Silver and Dunlap (1987), and the study was included. If multiple correlations were given in a study – for example using year marks for the first year, year marks for the second year and year marks total for separate correlations – the correlations with the broadest criterion were included (i.e. year marks total in this case). All included studies were assessed for independence to avoid including multiple correlations drawn from the same samples. On the other hand, one study (Noftle & Robins, 2007) had two independent samples, where the Big
Five personality traits had been correlated with GPA separately for each sample, and correlations for both of these samples were included due to their independence.

Following the guidelines outlined in the PRISMA Statement (Liberati et al., 2009) a flow diagram is presented in Figure 1 to illustrate the process of study selection. As a result of this process 20 individual empirical studies published in a variety of different journals and with a total of 21 independent samples were included for meta-analysis.

2.3. Data extraction process

Country of origin, study design, academic major and level of participants, sample size, mean age, distribution of gender, criterion, personality measure and predictors, and zero-order correlations were recorded for each of the included studies. The zero-order correlations were extracted by two researchers independently and compared for interrater reliability. There were no differences in ratings.

2.4. Meta-analytic procedures

The same meta-analytic procedures were undertaken for GPA-correlations with each of the Big Five personality traits in five separate analyses each following the approach of Borenstein, Hedges, Higgins, and Rothstein (2009) using Fisher’s $z$ transformations of correlations and utilising the software, CMA Version 2. In instances where sample sizes varied within a study without explicit information about how a conservative approach was chosen using the smallest sample size reported for the meta-analytic calculations. The random-effects model was chosen for computations to balance the relative weights assigned to the studies. The choice of the random-effects model was based on the fact that it did not seem reasonable to assume
Figure 1. Study selection process

doi:10.1371/journal.pmed1000097.
functional equivalence of the included studies. The studies have been carried out in
different countries with different university policies and selection procedures. And
while some studies had GPA – encompassing coursework as well as exams – as
criterion variable, others had overall exam marks exclusively. Even more importantly,
the samples are heterogeneous in that some consisted of psychology students only
while others included students from a variety of academic fields. This fact is exactly
what prompts an analysis of the potential moderator, academic major.
Notwithstanding the relatively conservative inclusion criteria applied to the selection
process earlier described it still would not be appropriate to assume a common, “true”
effect size shared by all the studies but rather similar effect sizes across studies. For
this reason the random-effects model was preferred over the fixed-effects model to
compute the summary effects for GPA-correlations with the Big Five personality
traits.

Subgroup analyses were performed for significant summary effects to assess
the impact of academic major as a potential moderator variable. Since samples
typically consisted of psychology students only or students from a variety of other
fields not specified, the moderator analyses were run with the two categories:
“psychology” versus “other”. The random-effects model was chosen again for this
purpose due to the diversity in studies given that behind the category “other” there
was a whole range of academic majors.

A problem faced by all researchers conducting meta-analyses as well as
literature reviews is the file drawer problem (Rosenthal, 1979) and the tendency for
studies with significant results and large effect sizes to find their way into the
published literature easier than studies with non-significant results (Dickersin, 1997).
To address this potential problem three different approaches dealing with publication
bias were applied. Rosenthal’s *Fail-safe N* was computed to determine the number of studies with small effects needed to be retrieved from “file drawers” to make the summary effects non-significant. Additionally, funnel plots with the effect sizes and standard errors of the studies were inspected to identify asymmetry and the possibility of publication bias. Finally, the Duval and Tweedie’s *Trim and Fill* method (Duval & Tweedie, 2000) was used to estimate the number of missing studies and re-compute the summary effects with these missing studies imputed into the model. The major advantage of this method is that it enables an estimation of the unbiased effect size and qualifies the conclusions reached in meta-analyses (Sutton, Duval, Tweedie, Abrams, & Jones, 2000). Of the four meta-analyses previously conducted (McAbee & Oswald, 2013; Poropat, 2009; Richardson et al., 2012; Trapmann et al., 2007) only one performed trim and fill procedures to handle publication bias (Richardson et al., 2012).

Furthermore, sensitivity analyses were conducted using the “one study removed” option in CMA in order to evaluate the impact of potential outliers on the summary effects.

### 3. Results

#### 3.1. Qualitative synthesis

From the 20 included studies with 21 independent samples 105 correlations were obtained for meta-analysis (21 correlations between each of the Big Five personality traits and GPA), and the aggregated sample size was 17,717. The extracted study characteristics are presented ordered by study author(s) in Table 1.

**Country:** The studies were primarily conducted in America and Great Britain.

**Design:** 11 of the studies were prospective while 9 were retrospective.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Country</th>
<th>Design</th>
<th>Academic major and level</th>
<th>N</th>
<th>A</th>
<th>F</th>
<th>Criterion</th>
<th>Personality measure and predictors</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chamorro-Premuzic, 2006</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 4 years of study</td>
<td>Psychology, undergraduate level</td>
<td>307</td>
<td>19.8</td>
<td>60.9</td>
<td>Overall exam marks, registry</td>
<td>NEO-FFI</td>
<td>.02</td>
</tr>
<tr>
<td>Chamorro-Premuzic &amp; Furnham, 2003a</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 3 years of study</td>
<td>Psychology, undergraduate level</td>
<td>70</td>
<td>19.8</td>
<td>70.0</td>
<td>Overall exam marks, registry</td>
<td>NEO-FFI</td>
<td>.22</td>
</tr>
<tr>
<td>Chamorro-Premuzic &amp; Furnham, 2003b</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 4 years of study</td>
<td>Psychology, undergraduate level</td>
<td>247</td>
<td>20.1</td>
<td>72.5</td>
<td>Overall exam marks, registry</td>
<td>NEO-PI-R</td>
<td>.07</td>
</tr>
<tr>
<td>Chamorro-Premuzic &amp; Furnham, 2008</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 1 year of study</td>
<td>Psychology, undergraduate level</td>
<td>137</td>
<td>19.2</td>
<td>70.0</td>
<td>Overall exam marks, registry</td>
<td>NEO-PI-R</td>
<td>.02</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Country</td>
<td>Design</td>
<td>Academic major and level</td>
<td>N</td>
<td>A</td>
<td>F</td>
<td>Criterion</td>
<td>Personality measure and predictors</td>
<td>r</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>---------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>--------------------------</td>
<td>------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Conard, 2006</td>
<td>USA</td>
<td>Retrospective design, academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Psychology, undergraduate level</td>
<td>289</td>
<td>19.5</td>
<td>77.7</td>
<td>GPA, self-report</td>
<td>NEO-FFI</td>
<td>.11</td>
</tr>
<tr>
<td>De Fruyt &amp; Mervielde, 1996</td>
<td>BEL</td>
<td>Prospective design, 6 months between administration of personality questionnaire and exams</td>
<td>Philosophy, history, languages, law, science, applied sc., economics, psychology and educational sc., applied biological sc., political and social sc., graduate level</td>
<td>741</td>
<td>23.4</td>
<td>53.6</td>
<td>Overall exam marks, registry</td>
<td>NEO-PI-R</td>
<td>.05</td>
</tr>
</tbody>
</table>

### Key

- **GPA**: Grade Point Average
- **NEO-FFI**: Five-Factor Personality Inventory
- **NEO-PI-R**: Revised Five-Factor Personality Inventory
- **r**: Correlation coefficient
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Country</th>
<th>Design</th>
<th>Academic major and level</th>
<th>N</th>
<th>A</th>
<th>F</th>
<th>Criterion</th>
<th>Personality measure and predictors</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dollinger, Matyja, &amp; Huber, 2008</td>
<td>USA</td>
<td>Retrospective design, academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Psychology (24%), and various other academic majors not specified, undergraduate level</td>
<td>338</td>
<td>21.9</td>
<td>61.0</td>
<td>Overall exam marks, registry</td>
<td>BFI</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Agreeableness</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>.26**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Extraversion</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Neuroticism</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Openness</td>
<td>-.08</td>
</tr>
<tr>
<td>Farsides &amp; Woodfield, 2003</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 3 years of study</td>
<td>Various academic majors not specified, undergraduate level</td>
<td>432</td>
<td>21.3</td>
<td>52.3</td>
<td>Overall exam marks, registry</td>
<td>NEO-FFI</td>
<td>.14**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Agreeableness</td>
<td>.14**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Extraversion</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Neuroticism</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Openness</td>
<td>.26**</td>
</tr>
<tr>
<td>Ferguson et al., 2000</td>
<td>GBR</td>
<td>Retrospective design, academic records obtained for 1 year of study</td>
<td>Medicine, undergraduate level</td>
<td>101-115</td>
<td>19.7</td>
<td>58.0</td>
<td>General medical training factor, registry</td>
<td>Goldberg’s Markers</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Agreeableness</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>.44**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Surgency</td>
<td>-.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Emotional Stability</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intellect</td>
<td>.01</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Country</td>
<td>Design</td>
<td>Academic major and level</td>
<td>N</td>
<td>A</td>
<td>F</td>
<td>Criterion</td>
<td>Personality measure and predictors</td>
<td>r</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------</td>
<td>-------------------------------</td>
<td>--------------------------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>------------------------------------</td>
<td>-----------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Furnham, 2012</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 1 year of study</td>
<td>Psychology, undergraduate level</td>
<td>178</td>
<td>19.4</td>
<td>80.9</td>
<td>Overall exam marks, registry</td>
<td>Agreeableness, Conscientiousness, Extraversion, Neuroticism, Openness</td>
<td>.13</td>
</tr>
<tr>
<td>Furnham et al., 2006</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 2 years of study</td>
<td>Psychology, undergraduate level</td>
<td>64</td>
<td>-</td>
<td>71.9</td>
<td>Overall exam marks, registry</td>
<td>Agreeableness, Conscientiousness, Extraversion, Neuroticism, Openness</td>
<td>.04</td>
</tr>
<tr>
<td>Gray &amp; Watson, 2002</td>
<td>USA</td>
<td>Retrospective design, academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Psychology, undergraduate level</td>
<td>334</td>
<td>19.2</td>
<td>63.8</td>
<td>GPA, self-report</td>
<td>Agreeableness, Conscientiousness, Extraversion, Neuroticism, Openness</td>
<td>.15**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.36**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.18**</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Country</td>
<td>Design</td>
<td>Academic major and level</td>
<td>N</td>
<td>A</td>
<td>F</td>
<td>Criterion</td>
<td>Personality measure and predictors</td>
<td>r</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------</td>
<td>----------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Komarraju, Karau, Schmeck, &amp; Avdic, 2011</td>
<td>USA</td>
<td>Retrospective design, academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Liberal arts, business, education, engineering, applied sciences/arts, mass communication, and agriculture, undergraduate level</td>
<td>308</td>
<td>-</td>
<td>52.3</td>
<td>GPA, self-reported</td>
<td>NEO-FFI</td>
<td>.22**</td>
</tr>
<tr>
<td>McKenzie et al., 2004</td>
<td>AUS</td>
<td>Prospective design, academic records obtained for 1 year of study</td>
<td>Various academic majors not specified, undergraduate level</td>
<td>1,193</td>
<td>21.4</td>
<td>50.5</td>
<td>GPA, registry</td>
<td>NEO-FFI</td>
<td>.08*</td>
</tr>
<tr>
<td>Musgrave-Marquart et al., 1997</td>
<td>USA</td>
<td>Retrospective design, academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Psychology, undergraduate level</td>
<td>161</td>
<td>20.8</td>
<td>64.0</td>
<td>GPA, registry</td>
<td>NEO-PI-R</td>
<td>.25**</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Country</td>
<td>Design</td>
<td>Academic major and level</td>
<td>N</td>
<td>A</td>
<td>F</td>
<td>Criterion</td>
<td>Personality measure and predictors</td>
<td>r</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>---------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-----------------</td>
<td>------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Nguyen, Allen, &amp; Fracastoro, 2005</td>
<td>USA</td>
<td>Retrospective design, academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Various academic majors not specified, undergraduate and graduate level</td>
<td>360</td>
<td>-</td>
<td>51.4</td>
<td>GPA, registry IPIP</td>
<td>Agreeableness Conscientiousness Extraversion Emotional Stability Intellect</td>
<td>.05</td>
</tr>
<tr>
<td>Noftle &amp; Robins, 2007</td>
<td>USA</td>
<td>Retrospective designs</td>
<td>Academic records obtained for various lengths of study depending on participating students’ progression</td>
<td></td>
<td></td>
<td></td>
<td>BFI</td>
<td>Agreeableness Conscientiousness Extraversion Neuroticism Openness</td>
<td>.03**</td>
</tr>
<tr>
<td>Sample 1</td>
<td>USA</td>
<td>Academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Psychology, undergraduate level</td>
<td>10,497</td>
<td>19.0</td>
<td>63.0</td>
<td>GPA, self-report</td>
<td>Agreeableness Conscientiousness Extraversion Neuroticism Openness</td>
<td>.06**</td>
</tr>
<tr>
<td>Sample 2</td>
<td>USA</td>
<td>Academic records obtained for 4 years of study</td>
<td>Various academic majors not specified, undergraduate level</td>
<td>475</td>
<td>18.0</td>
<td>56.0</td>
<td>NEO-FFI</td>
<td>Agreeableness Conscientiousness Extraversion Neuroticism Openness</td>
<td>.13**</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Country</td>
<td>Design</td>
<td>Academic major and level</td>
<td>N</td>
<td>A</td>
<td>F</td>
<td>Criterion</td>
<td>Personality measure and predictors</td>
<td>r</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------</td>
<td>-------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>----</td>
<td>-------</td>
<td>------</td>
<td>------------------</td>
<td>-----------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Richardson &amp; Abraham, 2009</td>
<td>GBR</td>
<td>Prospective design, 9 months between administration of personality questionnaire and exam</td>
<td>Life sciences, humanities, social and cultural studies, undergraduate level</td>
<td>614</td>
<td>20.3</td>
<td>76.9</td>
<td>GPA, registry</td>
<td>Agreeableness</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>Extraversion</td>
<td>.27**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Neuroticism</td>
<td>Openness</td>
<td>-.09</td>
</tr>
<tr>
<td>Saklofske et al., 2012</td>
<td>GBR</td>
<td>Prospective design, academic records obtained for 1 year of study</td>
<td>Psychology, undergraduate level</td>
<td>152-163</td>
<td>20.0</td>
<td>77.7</td>
<td>Overall exam marks, registry</td>
<td>Agreeableness</td>
<td>-.18*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>Extraversion</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Neuroticism</td>
<td>Emotional Stability</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Openness/Intellect/Imagination</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Steele-Johnson &amp; Leas, 2013</td>
<td>USA</td>
<td>Retrospective design. Academic records obtained for various lengths of study depending on participating students’ progression</td>
<td>Various academic majors not specified, undergraduate and graduate level</td>
<td>719</td>
<td>19.8</td>
<td>69.0</td>
<td>GPA, self-report</td>
<td>Agreeableness</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conscientiousness</td>
<td>Extraversion</td>
<td>.15**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Neuroticism</td>
<td>Openness</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.16**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>

Notes: \( N \) = overall sample size; \( A \) = sample mean age (years); \( F \) = percentage of female students in the sample; \( r \) = zero-order correlations.

* \( p < .05 \)  ** \( p < .01 \)  *** \( p < .001 \)
Academic major and level: 11 studies had samples with psychology students only, and 17 of the studies had samples drawn exclusively from the undergraduate level. 

N: Sample sizes for the individual studies ranged from 64 to 10,497. For two of the studies the sample size in the table is an interval as it varied in the correlational analyses for the Big Five personality traits. 

A: Reported mean ages were ranging from 18.0 to 23.4. Three studies failed to report mean age. 

F: Percentage of females in the sample was not available in all studies. For studies not reporting percentages the percentage of females was calculated whenever possible. The reported and calculated percentages of females ranged from 50.5% to 80.9% with large female percentages especially in samples with psychology students. 

Criterion: To stay true to the original studies it is reported whether European overall exam marks or American GPA was recorded, but as mentioned earlier the two constructs are treated as one in the statistical analyses and both denounced GPA in the rest of this paper. Half of the studies had overall exam marks as criterion, and half of them had GPA – mostly registry based. 

Personality measure and predictors: The nomenclature for the Big Five personality traits varies in the personality measures for three of the factors. This diversity implies that for example the factor in Goldberg’s Markers (1992) corresponding to Extraversion in the NEO-PI-R is denoted Surgency whereas the Openness factor in NEO-PI-R has a corresponding factor named Intellect. These differences originate in the historical development of the various Big Five personality measures (John et al., 2008), and this is outside the scope of the current meta-analysis. For the factor, Emotional Stability, corresponding to Neuroticism in the NEO-PI-R there is a practical consequence of the diversity though: Emotional Stability is a reversed
version of Neuroticism. The correlations between Emotional Stability and GPA outlined in Table 1 were therefore reversed in the meta-analytic calculations to ensure comparability.

\( r: \) Correlations for Conscientiousness and GPA were consistently positive and generally small–medium following Cohen’s guidelines (Cohen, 1977). For Agreeableness correlations were positive and small, often non-significant. Correlations for Neuroticism were mostly non-significant and alternately positive and negative. Likewise, for Openness correlations were most often non-significant and in both directions with an overweight of positive correlations. Extraversion yielded non-significant correlations only; some positive others negative.

3.2. Meta-analysis

Meta-analytic results are presented in Table 2. For each of the Big Five personality traits the weighted summary effect is reported together with the corresponding 95% confidence interval, the \( I^2 \) statistic\(^2\) and \( Q \) statistic\(^3\), aggregated sample size and number of independent correlation coefficients on which each weighted summary effect was based. Additionally, the unweighted summary effect, the standard deviation, the minimum and the maximum zero-order correlation is reported. Rosenthal’s \textit{Fail-safe N} is stated estimating the number of missing studies needed to nullify the weighted summary effect (i.e. \( p > .05 \)). Finally, based on the weighted summary effect and the funnel plot an estimation of the number of missing studies due to publication bias is reported, and if this number exceeds 0, a re-

\(^2\) For more on this see Higgins & Thompson (2002).

\(^3\) For more on this see Cochran (1954).
<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>k</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>r</th>
<th>r⁺</th>
<th>95% CI⁺</th>
<th>I²</th>
<th>Q</th>
<th>Fail-safe N</th>
<th>ka</th>
<th>r⁺ᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreeableness</td>
<td>17,717</td>
<td>21</td>
<td>.09</td>
<td>-.18</td>
<td>.25</td>
<td>.09***</td>
<td>.08***</td>
<td>[0.05, 0.11]</td>
<td>54.50%</td>
<td>43.96***</td>
<td>280</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>17,717</td>
<td>21</td>
<td>.10</td>
<td>.09</td>
<td>.44</td>
<td>.28***</td>
<td>.26***</td>
<td>[0.23, 0.30]</td>
<td>72.10%</td>
<td>71.67***</td>
<td>3,625</td>
<td>4</td>
<td>.24***</td>
</tr>
<tr>
<td>Extraversion</td>
<td>17,717</td>
<td>21</td>
<td>.07</td>
<td>-.11</td>
<td>-.16</td>
<td>-.00</td>
<td>-.00</td>
<td>[-0.03, 0.02]</td>
<td>29.08%</td>
<td>28.20</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>17,717</td>
<td>21</td>
<td>.12</td>
<td>-.35</td>
<td>.22</td>
<td>-.02**</td>
<td>-.01</td>
<td>[-0.05, 0.03]</td>
<td>73.64%</td>
<td>75.86***</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Openness</td>
<td>17,717</td>
<td>21</td>
<td>.11</td>
<td>-.15</td>
<td>.26</td>
<td>.06***</td>
<td>.07***</td>
<td>[0.03, 0.11]</td>
<td>71.18%</td>
<td>69.39***</td>
<td>222</td>
<td>0</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

N = aggregated sample size. k = number of independent samples. SD = standard deviation. Min. = minimum. Max. = maximum. r = unweighted summary effect. r⁺ = weighted summary effect. CI⁺ = confidence interval for weighted summary effect. I² = Higgins and Thompson’s measure of heterogeneity. Q = Cochrane’s measure of homogeneity. ka = number of missing studies imputed. r⁺ᵇ = weighted summary effect with missing studies imputed using Duval and Tweedie’s trim and fill.

** p < .01. *** p < .001.
computed, weighted summary effect with these missing studies imputed applying Duval and Tweedie’s trim and fill procedure is reported also. Inspection of the funnel plots revealed that publication bias was present in one instance, namely in the analysis of Conscientiousness-GPA correlations, and a re-computed weighted summary effect is reported accordingly. The funnel plot for Conscientiousness with missing studies imputed is shown in Figure A1 in the Appendix.

Sensitivity analyses employing the “one study removed” option in CMA did not reveal any significant impact of any one study’s correlation on the summary effect for any of the five meta-analyses. For this reason additional meta-analyses excluding individual outlier studies were not performed.

Conscientiousness was by far the best predictor of GPA with a weighted summary effect of .26 before trim and fill and .24 after. This very modest lowering of the summary effect after the trim and fill procedure together with the extremely high value of Rosenthal’s *Fail-safe N* shows the robustness of Conscientiousness in predicting tertiary GPA. Agreeableness and Openness were weak though statistically significant positive predictors of GPA with weighted summary effects of .08 and .07 respectively. Neuroticism and Extraversion both were insignificant predictors of GPA.

Inspection of forest plots revealed lack of overlap in confidence intervals for all traits indicating heterogeneity in effects. For the three significant weighted summary effects involving the traits Conscientiousness, Agreeableness, and Openness the $Q$ statistic was significant suggesting true heterogeneity of effect sizes in the 20 included studies not due to within-study error. Similarly, in all three cases $I^2$ yielded a
moderate-high proportion of real, non-spurious variance following guidelines provided by Higgins, Thompson, Deeks, and Altman (2003).

Some of this heterogeneity might stem from moderators. Subgroup analysis was performed for each of the three personality traits with significant weighted summary effects to assess the impact of the potential moderator: academic major. For these three traits Table 3 gives the weighted summary effect with its associated 95% confidence interval and $Q$ statistic for the two subgroups, “psychology” and “other”, together with the aggregated sample size for the subgroups, the number of independent samples on which each weighted summary effect was based, the unweighted summary effect, the standard deviation, the minimum and maximum zero-order correlation, and the proportion of between-studies variance explained by the moderator, academic major. For Openness and Agreeableness the weighted summary effects for the two subgroups, “psychology” versus “other”, did not differ substantially from each other. For Conscientiousness, however, there was a considerable difference between the weighted summary effects with a markedly stronger association of Conscientiousness with GPA for psychology students; $r = .31$ than for students from other academic fields; $r = .22$. The forest plot for the relationship between Conscientiousness and GPA in the included studies when divided into the two subgroups is shown in Figure A2 in the Appendix. For all three traits a large proportion of the between-studies variance could be explained by the moderator, academic major ($R^2 = .68 - .89\%$).

3. Discussion

3.1. Comparison of main results to previous research
Table 3. Results from subgroup analyses

<table>
<thead>
<tr>
<th>Academic major</th>
<th>N</th>
<th>k</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>r</th>
<th>r²</th>
<th>R²</th>
<th>CIr+</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreeableness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychology</td>
<td>12,436</td>
<td>11</td>
<td>.12</td>
<td>-.18</td>
<td>.25</td>
<td>.08***</td>
<td>.07*</td>
<td>.89</td>
<td>[0.02, 0.12]</td>
<td>.41</td>
</tr>
<tr>
<td>Other</td>
<td>5,281</td>
<td>10</td>
<td>.07</td>
<td>.02</td>
<td>.22</td>
<td>.09***</td>
<td>.09***</td>
<td></td>
<td>[0.05, 0.13]</td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychology</td>
<td>12,436</td>
<td>11</td>
<td>.09</td>
<td>.14</td>
<td>.42</td>
<td>.32***</td>
<td>.31***</td>
<td>.68</td>
<td>[0.26, 0.37]</td>
<td>5.05*</td>
</tr>
<tr>
<td>Other</td>
<td>5,281</td>
<td>10</td>
<td>.10</td>
<td>.09</td>
<td>.44</td>
<td>.23***</td>
<td>.22***</td>
<td></td>
<td>[0.17, 0.27]</td>
<td></td>
</tr>
<tr>
<td>Openness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychology</td>
<td>12,436</td>
<td>11</td>
<td>.11</td>
<td>-.15</td>
<td>.24</td>
<td>.07***</td>
<td>.09**</td>
<td>.68</td>
<td>[0.02, 0.15]</td>
<td>.35</td>
</tr>
<tr>
<td>Other</td>
<td>5,281</td>
<td>10</td>
<td>.10</td>
<td>-.09</td>
<td>.26</td>
<td>.06***</td>
<td>.06</td>
<td></td>
<td>[0.00, 0.12]</td>
<td></td>
</tr>
</tbody>
</table>

Notes: N = aggregated sample size. k = number of independent samples. SD = standard deviation. Min. = minimum. Max. = maximum. \( r \) = unweighted summary effect. \( r^2 \) = weighted summary effect. \( R^2 \) = proportion of between-studies variance explained by the moderator, academic major. CIr+ = confidence interval for weighted summary effect. \( Q \) = Cochran’s measure of homogeneity.

* \( p < .05 \). ** \( p < .01 \). *** \( p < .001 \).
The first objective of the current meta-analysis was to find out whether the prediction of academic performance by the Big Five personality traits with restricted inclusion criteria pertaining to personality measures as well as to academic performance would exceed results from previously conducted meta-analyses with broader criteria (McAbee & Oswald, 2013; Poropat, 2009; Richardson et al., 2012; Trapmann et al., 2007). The summary effect of .26 for the association of Conscientiousness with GPA in the current meta-analysis approximates the highest correlation for Conscientiousness found in the previous meta-analyses: .27 from Trapmann et al. (2007). In the current meta-analysis, however, no statistical corrections were made contrary to the psychometric meta-analysis conducted by Trapmann et al. As pointed out by Borenstein et al. (2009) the average effect size corrected for the effects of artifacts will generally be larger than if the corrections were not made. In line with this the uncorrected effect size also found in the meta-analysis by Trapmann et al. indeed was markedly lower; .21. For this reason the “raw” summary effect for Conscientiousness and GPA found in the current meta-analysis actually exceeds previous meta-analytic findings.

For the remaining Big Five personality traits also the correlation pattern was similar to the results from the four previously conducted meta-analyses with only vague correlations with GPA; Agreeableness and Openness being the strongest predictors of GPA apart from Conscientiousness.

3.2. The contribution of the moderator, academic major

The overall finding of Conscientiousness being a robust predictor of GPA with the relatively narrow operationalization of predictors and criterion in the current meta-analysis is not in itself surprising given the results from previous, broader met-
analyses. It is of course noticeable that at the tertiary level of education where range restriction in cognitive abilities diminishes their predictive validity (Jensen, 1980) Conscientiousness seemingly rivals psychometric intelligence in predicting GPA. But the more important contribution of the current meta-analysis pertains to the second objective: to find out whether academic major moderated the relationship between the Big Five personality traits and GPA. The subgroup analyses revealed differences between psychology students and students from other academic fields regarding the association of Agreeableness, Conscientiousness, and Openness with GPA; Conscientiousness especially. The association between Conscientiousness and GPA was markedly stronger for the psychology students than for students from other academic fields. To get a grasp of the magnitude of the association of the moderator with the effect sizes it is instructive to look at the calculated $R^2$. In primary studies it is custom to report $R^2$ to describe the impact of a covariate where $R^2$ is defined as the proportion of total variance explained by the covariate. In meta-analytic subgroup analysis it is possible to calculate an $R^2$ defined slightly differently: the proportion of true variance (between-study variance) explained by a covariate. The values in Table 3 for $R^2$ are such calculated proportions of the true variance between studies in trait-GPA correlations explained by the moderator, academic major. The fact that large proportions of the relatively small differences in summary effects between the subgroups for the traits, Agreeableness and Openness, could be explained by subgroup membership, 89% and 68%, respectively, is of course an interesting finding. But more importantly, the substantial difference in summary effects between the subgroups for Conscientiousness also could be explained to a fairly high extent by group membership ($R^2 = 68\%$). This implies that academic major was indeed an

---

4 For more on this see Borenstein, Hedges, Higgins, and Rothstein (2009).
important moderator of the Conscientiousness-GPA relationship, and this is a new
finding with high relevance for researchers investigating the personality-academic
performance interface.

Typically, studies of relationships between personality and academic
performance in tertiary education use samples consisting of psychology students only
since many universities systematically perform various psychological tests as an
obligatory part of enrolment in psychology courses. In the current meta-analysis this
sampling bias was evident in that half of the included studies had samples consisting
of psychology students only. But the results are most often generalized to students in
tertiary education overall – not just to psychology students. The findings in the
current meta-analysis suggest that these generalizations are not entirely valid and that
it is necessary to look at the academic fields of students when gauging the predictive
validity of the Big Five personality traits for academic performance. Roughly
speaking, we know a lot about psychology students and little about the remaining
population as pointed out by others (e.g. Cooper, McCord, & Socha, 2011; Reynolds,
2010). A related problem is that very often it is not specified in these studies from
which academic majors the participants came if not from psychology. On the one
hand, this limited the opportunity in the current meta-analysis to gauge the impact of
the moderator, academic major, to comparing only two groups: psychology students
versus “others”. On the other hand, this fact highlights the need to start taking
academic major into account in this kind of research. It would be very interesting
indeed to investigate the moderating effect of academic major on the
Conscientiousness-GPA association across multiple academic majors. Maybe some
alternative psychological predictors have more predictive validity and
Conscientiousness less in some non-psychology academic fields but not in others.
More studies with well-defined samples from other majors than psychology are needed in order to explore these questions. It also is probable that the remaining four personality traits have differential predictive validity depending on academic major just as they have been shown to have for job performance in specific occupations (Barrick & Mount, 1991). The significant difference between psychology students and students from other academic fields pertaining to the magnitude of the predictive validity of Conscientiousness found in this meta-analysis suggests that such research with broader sampling would be fruitful if undertaken.

3.3. Dropout

Another comment on the samples is the unknown factor of dropout in the studies used for meta-analysis. It is not custom in this research field to report how dropout was related to the effects found. For longitudinal studies in academic performance there will inevitably be some percentage of dropout, and if this dropout group differs substantially in personality from the group finishing a academic degree, this should be taken into account when interpreting results from this kind of studies. Robbins et al. (2004) performed a broad meta-analysis with multiple psychological and study skill factors as predictors of retention and GPA, and they found moderate relationships between retention and academic goals, academic self-efficacy, and academic-related skills. These findings suggest that goal-directed students with feelings of competence and ability to structure and manage their study habits are more likely to finish their studies. How the personality trait Conscientiousness specifically relates to retention was not evident from this meta-analysis though. A very recent study in retention at university sheds some light on this matter. Alarcon and Edwards (2013) explored several predictors of retention, and Conscientiousness indeed turned
out as a significant positive predictor. This finding is of interest for universities and public in general since high retention rates are desirable both ideologically and economically.

For interpretation of the results from the current meta-analysis the association between Conscientiousness and retention also is relevant. If retention is positively predicted by Conscientiousness, the group finishing a university degree will have some personality range restriction pertaining to this trait. And when correlations are calculated based on this selected sample, the “true” correlation without this range restriction could in fact be larger, similar to what is the case with intelligence and academic performance (Jensen, 1980). This only adds to the credibility of Conscientiousness as a predictor of academic performance.

3.4. Alternative approaches: facets, trait-complexes and curvilinear relationships

Wittmann and Süß (1999) have pointed out that problems in prediction of psychological outcomes often arise due to the lack of correspondence in breadth of predictors and criteria. In order to maximise prediction, Wittmann and Süß stress the notion of symmetry in aggregation and level of generality for predictor and criterion – also known as Brunswik symmetry. Could there be problems in symmetry between the Big Five personality traits and GPA, the pair of currently used variables?

The integrative trait-complex approach by Ackerman and Heggestad (1997) encompassing cognitive, affective, and motivational variables is very inclusive and has indeed proven useful in predicting both grades and attrition rates in tertiary education (Ackerman, Kanfer, & Beier, 2013). This broader approach, compared to the Big Five, rests on the assumption that personality traits alone are too narrow predictors of GPA, and this is an example of dealing with Brunswik symmetry.
Conversely, it may be that the Big Five personality traits are not too narrow but rather too broad to be effective predictors. Each of the Big Five personality traits comprises 6 subordinate facets, and though these facets are all correlated with their respective factor, they are largely independent from each other (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996) and therefore not likely equally correlated with other variables. In the domain of job performance facets of Conscientiousness have shown incremental validity in predicting job performance above that predicted by their superordinate factor (Ones & Viswesvaran, 2011). Something similar is likely the case in the domain of academic performance, but few studies have explored this. Of the 20 studies included in the current meta-analysis only three explored this possibility (Chamorro-Premuzic & Furnham, 2003b; De Fruyt & Mervielde, 1996; Gray & Watson, 2002). As hypothesized, some facets subordinate to Conscientiousness showed superior predictive validity in two of these studies (Chamorro-Premuzic & Furnham, 2003b; Gray & Watson, 2002), namely *Achievement striving*, *Dutifulness*, and *Self-discipline*. Interestingly, also a few facets subordinate to the factors Neuroticism and Extraversion were far better (negative) predictors of GPA than their corresponding factors. These relatively rare instances of research at the facet level are reviewed in depth elsewhere (O'Connor & Paunonen, 2007).

A related issue is the general assumption of linearity that dominates the field of statistics in psychological research. Our models depend on it, and very few test for curvilinear relationships between variables. It is indeed plausible that an excessive amount of a trait or facet normally considered beneficial in relation to e.g. academic success, such as Conscientiousness, can be impeding. A student extremely high on Conscientiousness could be too focused on details and on fulfilling responsibility and
might lose perspective of the study process and burn out eventually. Similarly, the seeming lack of association between Neuroticism and GPA could be due to a curvilinear relationship between these variables where some amount of Neuroticism could be beneficial and motivating whereas too much would cause adverse effects such as worrying and stress. None of the studies eligible for inclusion in the current meta-analysis tested whether there could be curvilinear relationships between the Big Five personality traits and GPA. However, Robbins, Allen, Casillas, Peterson, and Le (2006) tested for curvilinear relationships between college GPA and both Emotional Control and Social Activity from the Student Readiness Inventory (Le, Casillas, Robbins, & Langley, 2005), and inverted U-shaped relationships with GPA were found for these constructs resembling Neuroticism and Extraversion, respectively. On the same note, Cucina and Vasilopoulos found an inverted U-shaped relationship between Conscientiousness and GPA and a U-shaped relationship between Openness and GPA (Cucina & Vasilopoulos, 2005). Future research should explore this further.

To recapitulate: Behind the summary effect for Conscientiousness of .26 there is a hidden pattern of diverse correlations at the facet level with some higher and some lower; probably involving curvilinearity too. And in relation to the four personality traits not as highly correlated with GPA as Conscientiousness there could potentially be specific facets with considerable predictive validity not uncovered at present due to the lack of studies exploring the facet level as well as curvilinear relationships.

### 3.5. Mechanisms behind the Conscientiousness-GPA relationship

Returning to the findings of the current meta-analysis, it is worthwhile exploring from where the Conscientiousness-GPA relationship moderated by
academic major stems. A couple of studies included in the current meta-analysis pursued the mechanisms by which Conscientiousness predicts GPA. Highly conscientious students tend to show up for class more often than less conscientious students (Chamorro-Premuzic & Furnham, 2003a; Conard, 2006) and also employ learning strategies to a greater extent (McKenzie et al., 2004). Whether this directly mediate the relationship between Conscientiousness and GPA is not clear from the studies due to correlational study designs, but this is a possibility. A construct actually found to be a direct mediator of this relationship is Achievement Motivation (Richardson & Abraham, 2009). This seems somewhat self-evident when keeping in mind the superior predictive validity of the Conscientiousness facet Achievement Striving previously described. But it highlights the will-component of Conscientiousness that facilitates performance in general and academic performance in particular. The complex mechanisms behind the association of Conscientiousness and GPA are not completely understood yet, but the first steps on the road to this have been taken.

4. Limitations

In continuation of the issues outlined in the discussion there are some limitations. The over-reliance on psychology students and convenience sampling in the studies limited the options in the subgroup analysis to comparing only two groups: psychology students versus “others”. This is not ideal given the heterogeneity of majors in the group, “others”.

The assumption of linearity in the studies included also pertains to the meta-analysis itself. Possible curvilinear relationships between the Big Five personality traits and GPA could not be tested for.
The current meta-analysis did not correct for range restriction in the Big Five personality traits. Since Conscientiousness has been shown to predict academic performance at lower levels of education also (Poropat, 2009), the subjects studied at the tertiary level may well be a selected group with a higher level of Conscientiousness than the general population. This would lead to range restriction in predictor and an underestimation of the derived meta-analytic summary effect in the current meta-analysis. The advantage of this approach is that the uncorrected correlation informs us about the association between Conscientiousness and GPA within the selected population of university students, which in many instances is specifically what we are interested in.

A more construct-related limitation refers to the question whether the observed relationship between Conscientiousness and GPA is completely mediated by Achievement Motivation as suggested by Richardson and Abraham (2009). Research is scarce on this topic, and the current meta-analysis could not test for this possibility. However, the fact that Conscientiousness facets different from Achievement striving, such as Dutifulness and Self-discipline, also have been found to have predictive validity, as reviewed in the discussion, does not support that claim.

5. Conclusion

The discipline of meta-analysis is restricted by nature to summarize findings of the research available. Further research pursuing aspects such as the predictive validity of facets and curvilinear relationships with a heightened focus on non-psychology students would indeed qualify and enhance our knowledge about the relation between the Big Five personality traits and GPA. What we can say with certainty by now is, that with a summary effect of .26 between Conscientiousness and
GPA, personality places itself as a highly relevant predictor of the same magnitude as cognitive abilities. Abilities are important, but more so is what we actually *do* with our abilities as suggested almost a century ago by Edward Webb.
Appendix

Figure A1. Funnel plot for Conscientiousness with missing studies imputed.

Note. Imputed, missing studies are shown as filled circles. Studies originally included are shown as open circles. Summary effect without missing studies imputed is shown as an open diamond. Re-computed summary effect with missing studies imputed is shown as a filled diamond.
Figure A2. Forest plot for the Conscientiousness-GPA relationship.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Study Name</th>
<th>Correlation Limit Lower</th>
<th>Correlation Limit Upper</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>De Fruyt &amp; Mervielde, 1996</td>
<td>0.000</td>
<td>0.272</td>
<td>0.345</td>
<td>7.815</td>
</tr>
<tr>
<td>Other</td>
<td>Dollinger, Mejia, &amp; Huber, 2008</td>
<td>0.000</td>
<td>0.156</td>
<td>0.317</td>
<td>4.811</td>
</tr>
<tr>
<td>Other</td>
<td>Fenson &amp; Woodfield, 2003</td>
<td>0.000</td>
<td>-0.004</td>
<td>0.193</td>
<td>1.069</td>
</tr>
<tr>
<td>Other</td>
<td>Kang &amp; Sznajd, 1999, 1995</td>
<td>0.000</td>
<td>0.386</td>
<td>0.515</td>
<td>8.994</td>
</tr>
<tr>
<td>Other</td>
<td>Komarjun, Rhee, Schuer, &amp; Au, 2011</td>
<td>0.000</td>
<td>0.184</td>
<td>0.398</td>
<td>2.314</td>
</tr>
<tr>
<td>Other</td>
<td>Mahrenkel, Gau, &amp; Schmerer, 2004</td>
<td>0.000</td>
<td>0.045</td>
<td>0.116</td>
<td>3.021</td>
</tr>
<tr>
<td>Other</td>
<td>Ng, Lin, &amp; Pracenski, 2018</td>
<td>0.000</td>
<td>0.112</td>
<td>0.275</td>
<td>4.193</td>
</tr>
<tr>
<td>Other</td>
<td>Akhras &amp; Rabin, 2017</td>
<td>0.000</td>
<td>0.106</td>
<td>0.340</td>
<td>3.384</td>
</tr>
<tr>
<td>Other</td>
<td>Steele-Johnson &amp; Lewis, 2013</td>
<td>0.000</td>
<td>0.078</td>
<td>0.231</td>
<td>4.344</td>
</tr>
<tr>
<td>Other</td>
<td>Other</td>
<td>0.221</td>
<td>0.173</td>
<td>0.267</td>
<td>8.902</td>
</tr>
<tr>
<td>Psychology</td>
<td>Chamorro-Premuzic, 2006</td>
<td>0.420</td>
<td>0.325</td>
<td>0.528</td>
<td>7.066</td>
</tr>
<tr>
<td>Psychology</td>
<td>Chamorro-Premuzic &amp; Furnham, 2003a</td>
<td>0.360</td>
<td>0.111</td>
<td>0.271</td>
<td>3.371</td>
</tr>
<tr>
<td>Psychology</td>
<td>Chamorro-Premuzic &amp; Furnham, 2003b</td>
<td>0.360</td>
<td>0.246</td>
<td>0.494</td>
<td>5.987</td>
</tr>
<tr>
<td>Psychology</td>
<td>Chamorro-Premuzic &amp; Furnham, 2006</td>
<td>0.370</td>
<td>0.259</td>
<td>0.486</td>
<td>5.987</td>
</tr>
<tr>
<td>Psychology</td>
<td>Conard, 2006</td>
<td>0.370</td>
<td>0.244</td>
<td>0.494</td>
<td>6.813</td>
</tr>
<tr>
<td>Psychology</td>
<td>Furnham, 2012</td>
<td>0.370</td>
<td>0.234</td>
<td>0.479</td>
<td>6.170</td>
</tr>
<tr>
<td>Psychology</td>
<td>Furnham, Zhang, &amp; Chamorro-Premuzic, 2006</td>
<td>0.410</td>
<td>0.171</td>
<td>0.339</td>
<td>3.031</td>
</tr>
<tr>
<td>Psychology</td>
<td>Gray &amp; Watson, 2002</td>
<td>0.350</td>
<td>0.383</td>
<td>0.489</td>
<td>6.987</td>
</tr>
<tr>
<td>Psychology</td>
<td>MacGregor, Skirrow, &amp; Kiley, 1997</td>
<td>0.320</td>
<td>0.174</td>
<td>0.462</td>
<td>4.150</td>
</tr>
<tr>
<td>Psychology</td>
<td>McBeath, 2007</td>
<td>0.320</td>
<td>0.252</td>
<td>0.529</td>
<td>22.011</td>
</tr>
<tr>
<td>Psychology</td>
<td>Salovey, Mischel, &amp; O'Brien, 2000</td>
<td>0.340</td>
<td>0.260</td>
<td>0.527</td>
<td>9.038</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.252</td>
<td>0.274</td>
<td>0.300</td>
<td>12.473</td>
</tr>
</tbody>
</table>

Conscientiousness - GPA
References

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


Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219-245.


  doi:http://dx.doi.org.ez.statsbiblioteket.dk:2048/10.1016/j.jclinepi.2009.06.006


