

Personality Psychology

# The Big Five Across Socioeconomic Status: Measurement Invariance, Relationships, and Age Trends

Bradley T. Hughes<sup>1</sup> <sup>a</sup>, Cory K. Costello<sup>2</sup> , Joshua Pearman<sup>3</sup> , Pooya Razavi<sup>1</sup> , Cianna Bedford-Petersen<sup>1</sup> , Rita M. Ludwig<sup>1</sup> , Sanjay Srivastava<sup>1</sup> 

<sup>1</sup> Department of Psychology, University of Oregon, Eugene, OR, US, <sup>2</sup> Department of Psychology, University of Michigan, Ann Arbor, MI, US, <sup>3</sup> Department of Psychology, University of Georgia, Athens, GA, US

Keywords: personality, socioeconomic status, big five, measurement invariance, personality age trends

<https://doi.org/10.1525/collabra.24431>

---

## Collabra: Psychology

Vol. 7, Issue 1, 2021

---

Associations between socioeconomic status (SES) and personality traits have important implications for theory and application. Progress in understanding these associations depends on valid measurement, unbiased estimation, and careful assessment of generalizability. In this registered report, we used data from AIID, a large online study, to address three basic questions about personality and SES. First, we evaluated the measurement invariance of a common measure of personality, the Big Five Inventory, across indicators of educational attainment, income, and occupational prestige. Fit indices showed some instances of detectable noninvariance, but with little practical impact on substantive results. Second, we estimated associations between SES and personality. Results showed that personality and SES were largely independent (most  $r$ s < .1), in contrast to predictions derived from several previous studies. Third, we tested whether age trends in personality were moderated by SES. Results did not support predictions from social investment theory, but they did suggest that age trends were largely generalizable across SES. We discuss the implications of these findings for developing and validating personality measures for use in diverse samples. We also discuss the implications for theories that propose that the Big Five are responsive to, or partially responsible for, people's economic and social conditions.

A growing body of evidence has observed a relationship between personality traits and socioeconomic status (SES). For example, conscientiousness positively predicts SES over and above intelligence (Heckman & Kautz, 2012) and neuroticism is negatively associated with income (Kajonius & Carlander, 2017). Moreover, these relationships are linked to important outcomes; for example, personality can explain ~20% of the increased all-mortality risk associated with being lower in SES (Chapman et al., 2010). Documenting these kinds of associations is an important precursor to developing theories about transactional pathways between personality and social contexts, including how social roles and experiences might shape personality development and vice versa.

Before proceeding with such theoretical work, however, it is important to address basic methodological and descriptive empirical questions about personality and SES. Personality measures are often developed in samples of college students, but in order to study the associations between personality and SES, the measures have to function equivalently in more socioeconomically diverse samples. Few studies report analyses of measurement invariance before

presenting substantive associations. Thus, efforts to explain associations between SES and personality may be misguided if those associations are not replicable and generalizable.

In this research we address three important questions in the study of personality and SES. First, how valid are comparisons of personality across different socioeconomic strata? We address this question by analyzing the measurement invariance of a widely used personality measure, the Big Five Inventory (John & Srivastava, 1999), across multiple SES indicators. Second, how do Big Five personality traits correlate with education, income, and occupational prestige? We investigate the associations between the Big Five and socioeconomic indicators. And third, how well do age trends in personality generalize across levels of socioeconomic status? We consider mean-level age trends in the Big Five that are often described as representing typical developmental pathways in adulthood.

---

<sup>a</sup> Correspondence concerning this article should be addressed to: Bradley T. Hughes, 1227 University St, Eugene, OR 97403; phone: (541) 346-4921; email: bhughes7@uoregon.edu

## Socioeconomic Status and Its Relevance to Personality

SES is a widely studied contextual phenomenon in psychology. Broadly speaking, SES refers to an individual's standing within the social hierarchies of a society. More specific definitions and approaches vary, with some researchers emphasizing economic hierarchies and others emphasizing social variables like prestige (Bradley & Corwyn, 2002). In practice, SES is frequently measured through one or more of three variables: educational attainment, income, and occupational prestige (Adler et al., 1994; Diemer et al., 2013; Saegert et al., 2006). In the present research we will consider all three of these variables, each of which has been shown to be associated with personality traits (Jonassaint et al., 2011).

Why is it important to study the relationship between SES and personality? SES is defined by differential access to economic and social resources that can both affect (Bradley & Corwyn, 2002; Conger et al., 2002; Jonassaint et al., 2011) and be affected by personality (Damian et al., 2015; Heckman & Kautz, 2012). For example, childhood SES predicts patterns in personality, such that growing up in a high SES home is associated with higher levels of Extraversion and Openness, whereas growing up in a low SES home is associated with high neuroticism and low conscientiousness (Jonassaint et al., 2011). Similarly, an individual's personality measured in high school can predict future socioeconomic success after accounting for their parent's SES (Damian et al., 2015). Researchers and policymakers are increasingly recognizing the potential for interventions on personality development to affect social and economic indicators and vice versa, making this a key area for societally relevant basic and translational research (Bleidorn et al., 2019).

### Measurement Invariance

Prior to investigating the associations between SES and personality, it is important to assess if personality, as measured, reflects the same construct (i.e., has the same structure) across levels of SES. Measurement invariance concerns the extent to which a measure works similarly across levels of some factor or group (e.g., across different cultures, age groups, gender, income and education, etc.). It is important to examine measurement invariance for two reasons. First, measurement invariance is an important aspect of structural validity, or the extent to which the hypothesized structure of a measure is reflected empirically in the data (Flake et al., 2017; John & Soto, 2007). Thus, establishing measurement invariance is a psychometric goal in its own right, providing evidence that a proposed measurement model is structurally valid. Second, establishing measurement invariance is a precondition for unambiguously interpreting differences in that measured construct across groups. For example, an observed relation between conscientiousness and educational attainment (e.g., a correlation between conscientiousness and years of educational attainment, or a mean difference in conscientiousness between high school graduates and college graduates), could either reflect a true relation between these constructs or differ-

ences in how people of varying educational backgrounds respond to the items of the scale used to measure conscientiousness. Only by establishing measurement invariance across levels of educational attainment could an analysis like this unambiguously reflect a relation between the constructs.

Measurement invariance can be tested with multiple-group confirmatory factor analysis (CFA) in which parameters are constrained to be equal across levels of a grouping variable (Gregorich, 2006). Each level of measurement invariance enables the unambiguous interpretation of certain comparisons. *Metric invariance* is established when factor loadings are equal across groups, and it enables comparisons of variances and covariances among latent factors (e.g., a correlation between latent conscientiousness and education). *Scalar invariance* is established when item intercepts are equal, and it enables comparisons of both latent and observed means (e.g., a mean difference between high school and college graduates). *Strict factorial invariance* is established when residual variances are equal across groups, and it enables comparisons of observed variances and covariances (e.g., a correlation between conscientiousness scale scores and education; Gregorich, 2006). Our planned analyses include differences in Big Five observed means and covariances (e.g., with age). Strict factorial invariance most fully supports these analyses, though this is often considered too strict a test of measurement quality (Nye & Drasgow, 2011), and so our approach will include a pragmatic assessment of its impact.

To our knowledge, measurement invariance of the Big Five across different categories of SES has not been tested before, with one exception, where conscientiousness was shown to be partially invariant across SES (Ludwig et al., 2019). However, there is a rich literature on testing the measurement invariance across other variables, including nations (Thalmayer & Saucier, 2017), US states (Gebauer et al., 2014), age groups (del Barrio et al., 2006), cohorts (Borghuis et al., 2017), genders (Marsh et al., 2010), and ethnicity (Schmitt et al., 2011). These studies provide mixed results in terms of the measurement invariance of the Big Five, with several studies demonstrating configural and metric, but not scalar (nor full) invariance across different groups (e.g., Vecchione et al., 2012). Furthermore, a number of studies found measurement invariance for only a subset of the traits (e.g., Borghuis et al., 2017), and others were suggestive of partial invariance (e.g., Marsh et al., 2010). Most directly relevant to the current study, prior work has demonstrated differences in acquiescent responding across levels of education attainment (one pillar of SES), which could lead to differences in item loadings, intercepts, and residuals (Rammstedt et al., 2010). These results point out the necessity of conducting a measurement invariance analysis of the Big Five across SES categories prior to conducting further analyses for which the interpretations are conditional on the quality and equivalence of the measurement (Borsboom, 2006).

### Associations Between Personality and Socioeconomic Status

Relationships between socioeconomic status and Big

Five traits have been previously reported across multiple studies. Positive associations have been observed between SES and agreeableness and emotional stability as well as openness and extraversion (Jonassaint et al., 2011). The bulk of the evidence to date, however, has centered on the association of SES with conscientiousness and neuroticism.

Multiple longitudinal studies have linked conscientiousness and its facets to two components of SES, educational achievement (Chamorro-Premuzic & Furnham, 2003; Nofhle & Robins, 2007) and occupational attainment (Judge et al., 1999; Shiner et al., 2003). Further, conscientiousness has been shown to moderate individual SES over the life course, such that higher conscientiousness played a larger role in predicting future SES for those from low-SES backgrounds (Damian et al., 2015). Proposed mechanisms for the observed positive trend between conscientiousness and SES include proclivity for long-term planning (Ludwig et al., 2019) and the selective encouragement of behaviors early in development that are thought to promote future professional success (Roberts et al., 2007).

Whereas conscientiousness has been associated with positive socioeconomic outcomes, studies of SES and neuroticism have reported the opposite trend. Lower SES individuals tend to score higher in neuroticism (Lahey, 2009), and lower childhood SES predicts higher neuroticism in adulthood (Jonassaint et al., 2011). It is hypothesized that the lack of resources intrinsic to being low SES causes this observed trend of increased anxiety and depression (Santiago et al., 2011), and that being relatively high in neuroticism among this socioeconomic group can further increase the risk of mood disorders (Jokela & Keltikangas-Järvinen, 2011). Finally, some evidence suggests that SES and neuroticism interact to increase rates of all-cause mortality among the relatively disadvantaged (Chapman et al., 2010).

The reviewed evidence presents two clear hypotheses relative to individual SES: that SES is positively associated with conscientiousness, and that it is negatively associated with neuroticism. The observed effect sizes were in the small-to-moderate ( $r = .10 - .30$ ) range for all correlations (though we note that these estimates are not based on a full meta-analysis). There is some evidence to suggest that agreeableness and extraversion will be positively associated with SES as well, though the relatively smaller body of evidence provides less confidence for these traits.

### Generalizability of Age Trends Across SES

In the past two decades, a consensus has emerged about how mean levels of Big Five traits change across the adult lifespan. Accordingly, mean levels of conscientiousness and agreeableness increase with age; neuroticism decreases, especially among women; and openness and extraversion change relatively little. This pattern has been documented in longitudinal studies (Specht et al., 2011); cross-sectional studies (Soto et al., 2011; Srivastava et al., 2003), and a meta-analysis (Roberts et al., 2006). The pattern has been described as a generalizable principle of development (Roberts et al., 2008). We propose to test how well this pattern generalizes across levels of SES.

Mean-level change findings have been interpreted through a variety of lenses by multiple researchers, with

different implications for generalizability. In their five-factor theory, McCrae & Costa (2008) proposed the intrinsic maturation hypothesis, which held that these patterns reflect endogenous development that is invariant across social environments (McCrae et al., 1999, 2000). Five-factor theory therefore implies that age trends in the Big Five should be invariant across socioeconomic strata.

By contrast, Roberts et al. (2008) have described mean-level change as reflecting the maturity principle: during adulthood, the average person's personality changes to support them becoming a more productive, steady, and prosocial member of society. A second principle, the social investment principle, proposes a mechanism that drives these changes. According to the social investment principle, changes in personality result from psychological commitments to social institutions, such as work, family, and community. Through these role commitments, people are exposed to new expectations and contingencies that shape their behavior. If socioeconomic status is related to these role commitments and their normative timing, then according to the social investment principle, the magnitude and possibly even direction of mean-level change will vary as a function of socioeconomic class.

How do these broad principles translate into specific predictions about SES and age trends in personality? In one test of the social investment principle, Bleidorn et al. (2013) studied how nation-level differences in the timing of social roles that are linked to SES (e.g., marriage, employment) were associated with age trends. Early normative timing of the completion of education (used as a proxy for entering the workforce) was associated with more of a decrease in neuroticism (a more negative slope) and more of an increase in conscientiousness (a more positive slope). Early normative timing of family roles, as indicated by teen marriage rates, teen birth rates, and mean age at first marriage, was associated with a more positive age slope for openness. These national differences were interpreted as reflecting the social investment processes that have been theorized to operate at an individual level.

The social role variables studied by Bleidorn et al. (2013) can be translated into predictions for SES indicators. Timing of the completion of education is closely related to educational attainment, a common status indicator, with higher SES reflecting later normative completion of education. Marriage and family roles are also associated with socioeconomic status. Higher educational attainment is associated with later normative age of marriage (Parker & Stepler, 2017; Wang, 2018) and lower family socioeconomic status is associated with greater teen pregnancy (Penman-Aguilar et al., 2013). Thus, applying the theoretical framework and findings of Bleidorn et al. (2013), we can generate a prediction that higher socioeconomic status will be positively associated with the age slope for neuroticism, negatively associated with the age slope for conscientiousness, and negatively associated with the age slope for openness.

### The Present Research

The goal of the present study is to better understand the relationship between SES and personality by: (H1) testing if the BFI (John & Srivastava, 1999), a widely used personality

measure, is measuring the same constructs across economic groups using income, education, occupational prestige, and occupational income as indicators of SES, (H2) estimating the associations between Big Five traits and SES, in an economically diverse sample, to test if they are consistent with previous findings, and (H3) testing if the effects of the same four indicators of SES on mean-level age trends in personality development are consistent with differences in normative role timing.

We tested the following specific hypotheses for H2: (a) that the indicators of SES are positively associated with conscientiousness, agreeableness, and extraversion, and (b) that the indicators of SES are negatively associated with neuroticism. Based on the previously reviewed literature predicted correlations in the small-to-moderate ( $r = .10 - .30$ ) range. In addition to examining these bivariate relationships, using multiple regression, we estimated the unique relationship between three of the SES indicators, income, education, and occupational prestige and each of the Big Five traits.

Based on normative role timing, we tested three specific hypotheses for H3: (a) that SES is positively associated with the age slope for neuroticism, (b) SES is negatively associated with the age slope for conscientiousness, and (c) SES is negatively associated with the age slope for openness. In cross-sectional data such as in the present study, mean-level differences can also reflect cohort effects (Schaie, 1965). Convergence between cross-sectional and longitudinal approaches, the latter of which are unaffected by cohort, has largely led personality psychologists to interpret differences as reflecting age effects (Roberts et al., 2006). However, we also considered the possibility of cohort main effects or cohort-by-SES interactions in interpreting the results.

## Method

### Participants and Procedure

Data for this study comes from the Attitudes, Identities, and Individual Differences study (AIID; Hussey et al., 2018). AIID was a large-scale ( $N \approx 300,000$ ) study that was run on the Project Implicit website between 2004 and 2007. AIID participants were asked to complete demographic measures, including indicating their occupation, income, and educational attainment. In a planned missingness design, each AIID participant also completed a randomly selected set of individual-difference items. Two of the subsets included items from the Big Five Inventory (BFI; John & Srivastava, 1999). As a result of this design, the effective  $N$  for the present analyses will be smaller than the full AIID dataset.

The research team that collected the AIID dataset split it into two subsets, an exploratory set (15% of the total sample) and a hold-out set. The hold-out dataset was embargoed and made available only to researchers who had a Stage 1 registered report accepted for publication. Unless noted otherwise, analyses reported in the Method section were conducted in the exploratory set prior to preregistration, and analyses reported in the Results section were conducted in the hold-out, confirmatory dataset after Stage 1 acceptance.

**Exclusion Criteria.** We initially screened participants from the AIID data and only included those who reside in the United States, completed one of the two BFI subsets (Agreeableness and Openness or, Extraversion, Conscientiousness, and Neuroticism), and did not select “unemployed” as their occupation. We excluded “unemployed” individuals from the analyses for several reasons. First, being unemployed is a transient state, and without further information it is impossible to differentiate between short-term versus long-term unemployment, voluntary versus involuntary unemployment versus retirement, etc. Second, it is difficult to interpret self-reported income for individuals who select “unemployed” (e.g., given ambiguity in the instructions, some participants may have significant part-time or investment income, whereas others may have no current income but report a salary from their last job). Third, for these reasons we did not believe it would be valid to derive occupational indices for the “unemployed” category, which would leave unemployed participants missing on those variables. For analyses including the income measure, we also excluded participants who indicated “don’t know” because for our analysis this response is equivalent to missing data. We will use all available data for each analysis (pairwise deletion), which will result in a different number of participants for each analysis.

For the measurement invariance analysis, we included participants who indicated “student” as their occupation as a reference category, since a substantial amount of the evidence for the Big Five Inventory’s validity comes from student samples (John & Srivastava, 1999). For the Big Five and SES relationship (H2) and the age trend analysis (H3), we excluded students and those under 25 in order to focus on the effects for individuals that are more likely to have completed their educations (making the “education” variable less transient and more indicative of social position), entered the workforce, and established an occupation.

### Measures

Following recent recommendations for SES measurement (Diemer et al., 2013), we used three indicators of SES: educational attainment, income, and occupational prestige. In addition, because of concerns about lack of fidelity in the AIID income measure, we calculated an additional indicator of SES, *occupational income*.

**Educational attainment.** AIID participants reported their educational attainment. Different versions of the AIID questionnaire used slightly different sets of categories, so some of the categories in the analyses reflect combinations of categories from different versions. In addition, the “Not a high school graduate” had too few participants after exclusions, so we combined it with “High school graduate” to create a “No college” category. The final education categories for the analyses were: “No college,” “Some college or associate’s degree,” “Bachelor’s degree,” and “Graduate degree or graduate education.” For correlation and regression analyses, these categories were assigned the numbers 1-4 respectively and entered into the models as a continuous variable.

**Income.** Participants self-reported their household income by selecting one of the following five categories: less

**Table 1. Descriptive Statistics for the AIID Confirmatory Data Big Five Traits**

Trait	N	Mean	SD	Cronbach's Alpha
Extraversion	2168	3.87	1.03	.87
Agreeableness	2220	4.38	0.80	.79
Conscientiousness	2168	4.47	0.84	.83
Neuroticism	2168	3.24	0.96	.83
Openness	2220	4.79	0.69	.79

Note: BFI responses were on 6-point scale (range 1 to 6).

than \$25,000; \$25,000 to \$49,999; \$50,000 to \$74,999; \$75,000 to \$149,999; greater than \$150,000; and “don’t know.” For correlation and regression analyses, “don’t know” responses were excluded and the remaining categories were assigned to the numbers 1-5, respectively, and entered into the models as a continuous variable.

**Occupational prestige.** Respondents were asked to indicate their occupation by selecting one of 98 job categories, including “student,” that most closely matched their current occupation, or could indicate that they were unemployed. These 98 categories were derived by combining similar job titles from the roughly 1,000 job titles in the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA) O\*NET occupational data, which is a government database that includes a list of job titles and associated characteristics.

To calculate the prestige scores for these 98 job categories, we began with the prestige scores for all of the O\*NET job titles, collected in a study of people’s ratings of the prestige of various occupations (Hughes et al., n.p.). The prestige score for each job title represents the average ratings of the prestige of that job title from a minimum of 150 unique judges. The occupational prestige scores for each of the 98 AIID job categories were calculated as the average prestige for the job titles contained within that category (scores for the 98 job categories are available here: <https://osf.io/wv59a/>). For example, if a participant selected “Operations Management,” their occupational prestige ratings were the average of the prestige ratings of the 20 job titles within this category (e.g., financial managers, purchasing managers, human resource managers).

**Occupational income.** The income categories in the AIID study appeared too broad to capture the effects of interest. For example, the range from \$25,000 - \$50,000 may capture both individuals who are living in poverty and those from lower-middle class homes. In addition, self-reported income is commonly misreported by participants (Epstein, 2006). To address these concerns, we calculated a continuous occupational income score using national average income data from O\*NET. Occupational income data was available for 93 of the 98 scores, which were calculated in the same manner as prestige, by averaging the yearly income for each of the job titles within the category (available here: <https://osf.io/wv59a/>). Because the AIID data was collected from 2004 to 2007, we used the O\*NET income data for 2007 to estimate the occupational income for each of the categories.

**Personality traits.** Big Five personality traits were assessed with items from the Big Five Inventory (BFI; John & Srivastava, 1999). As part of the planned missingness design of the AIID study, participants did not complete the entire BFI. Instead, a subset of AIID participants were randomly assigned either one of two subsets of BFI items: one subset contained the Agreeableness and Openness items; the other subset contained Extraversion, Conscientiousness, and Neuroticism items. Responses were on a 6-point scale, with 1 anchored with “strongly disagree” and 6 anchored with “strongly agree.” The reversed keyed items in the BFI were reverse scored prior to distribution of the AIID data. Therefore, to calculate traits scores, we averaged the items for that trait. Reliability estimates for the confirmatory data are reported in [Table 1](#).

## Analysis Plan

### Inference Criteria

For H1 analyses, our interpretations and decisions were based on fit indices; we discuss these considerations in greater detail below. For H2 and H3 analyses, we calculated effect sizes (correlation and regression coefficients) and 95% confidence intervals, and based our interpretations on those statistics. In addition, we calculated *p*-values. Following the recommendations of Benjamin et al. (2018), we interpret *p* < .005 as “significant” and *p* < .05 as “suggestive.” Given the large number of hypotheses to be tested, we adopted this conservative approach to interpreting *p*-values reduce the chance of over-interpreting unreplicable effects.

### Measurement Invariance (H1)

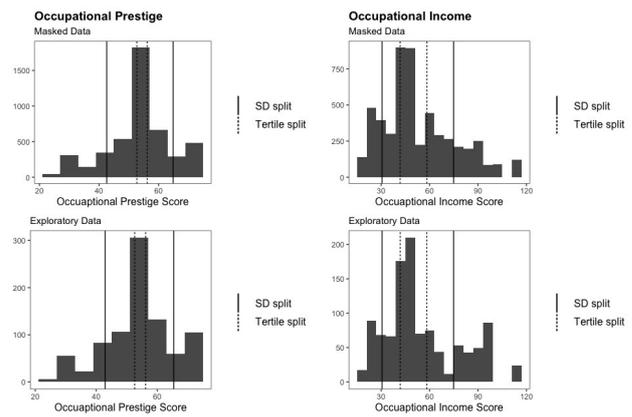
We examined measurement invariance across educational attainment, income, occupational income, and occupational prestige separately for each of the Big Five domains, resulting in a total of 20 sets of measurement invariance analyses. The correlation between occupational prestige and occupational income was substantially larger ( $r = .75$ ) than between any of the other indicators of SES (next-largest  $r = .45$ ). This relationship may be artificially inflated because both are based on the same aggregated occupation codes, but income and prestige are conceptually distinct. For example, there are prestigious occupations (e.g., teachers) that have low income and there are low prestige jobs (e.g., sanitation workers) that have high incomes. We acknowledge that the measurement invariance analyses conducted with each of these occupational indicators of SES are

unlikely to produce different results. However, because it is common practice for researchers to select a single indicator of SES for analysis it is important to assess invariance across both occupational measures.

Testing for measurement invariance across the continuous SES indicators required splitting the respondents into discrete groups (Gregorich, 2006). The educational attainment and income measures collected in the AIID study each had five response choices. The educational attainment responses included a small number of participants in both the “not a high school graduate” and “high school graduate” categories. We combined these categories into a “no college” category and tested for invariance across this category and the three remaining responses. The income measure had an adequate number of participants in each response choice so we considered each response choice a group and test for invariance across them. Unlike these “naturally” occurring response choice groups, the occupational measures are continuous and normally distributed (see Figure 1). We examined the distributions of occupational prestige and occupational income in both the exploratory and masked data and considered dividing them in a number of different ways, including tertile splits and by standard deviation. Figure 1 shows that dividing the data (both the exploratory and masked) into even tertiles results in individuals who clearly belong in the middle group being assigned to both the high and low groups. Dividing participants one standard deviation above and below the mean, instead of splitting the data into even groups, addressed this issue without relying on the idiosyncrasies of this data to determine groups. Based on these distributions, we split the occupational variables, prestige and income, into three groups with low, medium, and high scores. Using standard deviations in this way resulted in the medium group having more participants, and the “low” and “high” groups being more extreme than they would be with an evenly spaced tertile split. In our judgment, informed in part by observing where these cutpoints fell in the data, this approach fits better with the everyday meaning of low and high standing in socioeconomic hierarchies.

Each of the Big Five was fit as a single latent variable with all of the corresponding items treated as indicators; for example, extraversion was modeled as a latent variable with the eight BFI extraversion items as its indicators. For each comparison, we first fit an unrestricted model that allowed each group to have its own measurement mode, assuming only configural invariance (i.e., that the items on a priori scale have a one-factor structure). We then examined metric, scalar, and strict factorial invariance by assessing the decrement in fit in increasingly constrained, nested models produced by setting the loadings, intercepts, and residual variances equal across groups respectively. We assessed the magnitude of invariance by calculating effect sizes created by Nye & Drasgow (2011), which provided item- and scale-level estimates of the impact of differences in item loadings and intercepts across groups. Models were fit using the lavaan package (version  $\geq 0.6.4$ ; Rosseel, 2012) in R (version  $\geq 3.6.1$ ; R Core Team, 2019) and we used the dmacs package (version  $\geq 0.1.0$ ; Dueber, 2019) to calculate the effect sizes.

We evaluated measurement invariance by comparing model fit indices and the effect sizes developed by Nye &



**Figure 1. Distribution of Occupational Prestige and Income Scores with Tertile and Standard Deviation Splits**

Note: Figure created with ggplot2 package (Wickham, 2016).

Drasgow (2011). For fit indices, we used change in McDonald’s Noncentrality Index ( $\Delta\text{MFI}$ ; as recommended by Kang et al., 2016), change in Root Mean Squared Error of Approximation ( $\Delta\text{RMSEA}$ ), and change in Akaike Information Criterion ( $\Delta\text{AIC}$ ); however, we based our interpretations most heavily on  $\Delta\text{MFI}$  if the different fit measures were in conflict. We also used the three effect sizes proposed by Nye & Drasgow (2011). The first is an item-level effect size called  $d_{\text{MACS}}$  that represents the extent of invariance in the loadings and intercepts (together) of each item in a scale on a metric designed to be similar to Cohen’s  $d$ ; the empirically-derived benchmarks for small, medium, and large effects for  $d_{\text{MACS}}$  are .20, .40, and .70 respectively (Nye et al., 2019). In addition to item-level  $d_{\text{MACS}}$ , there are two scale-level estimates that capture the practical consequences of measurement invariance on observed scale scores. They are  $\Delta_{\text{Mean}}$  and  $\Delta_{\text{Var}}$  which correspond to expected differences in observed means and variances (respectively) due to measurement invariance (Nye & Drasgow, 2011; see also Clark et al., 2016).

Interpretation can be based on fit indices in a continuous and calibrated way, and several authors caution against using hard thresholds when not necessary (Kang et al., 2016; Nye & Drasgow, 2011). However, thresholds are often needed for behavioral decisions; one such decision is whether to proceed with analyses using an *a priori* scale, or to drop differentially functioning items in order to produce an invariant scale. Regardless of the results of the measurement invariance analyses, we conducted the remaining analyses (for H2 and H3) with a priori Big Five scale scores. In addition, for Big Five scale(s) that showed a  $\Delta\text{MFI}$  value greater than .01 from configural to metric invariance or metric to scalar invariance and a  $\Delta_{\text{Mean}}$  greater than .20, we constructed partially invariant *reduced observed scales* based on a subset of items, guided principally by items’  $d_{\text{MACS}}$  (i.e., by dropping items with high  $d_{\text{MACS}}$ ). We report results for the invariant scale(s) alongside the *a priori* scales and calibrate our interpretations accordingly. The choice of a  $\Delta\text{MFI}$  value of .01 as a cutoff was based on the recommendation of Kang et al. (2016) and the choice of a  $\Delta_{\text{Mean}}$  of .20 was based on practical considerations (i.e., that a difference

of at least 0.20 on a scale mean is non-negligible).

**Selecting the reference indicator.** When pursuing a partial invariance model, the choice of the reference indicator becomes consequential, as using a noninvariant reference indicator might change the outcome of the measurement invariance analyses (Yoon & Millsap, 2007). Based on simulation studies investigating measurement invariance across two groups, with balanced sample sizes, Jung & Yoon (2017) recommend choosing the item with the smallest modification index as the reference indicator. Although the present research will test for invariance across more than two groups, with uneven group sizes, which makes Jung and Yoon's (2017) recommendations not directly applicable, the logic behind choosing the item with the smallest modification index as the reference indicator is reasonable. Accordingly, when we found evidence of noninvariance or partial invariance for any of the models, we used the following steps to select the reference indicator: 1) run a full invariance model, 2) calculate the mean of modification indices for each indicator, and 3) select the indicator with the smallest mean of modification indices as the reference indicator.

In addition to testing measurement invariance, we looked at the structural invariance of the Big Five by estimating the extent to which factor correlations vary across groups. As noted above, participants completed one of two subsets of Big Five items, meaning we can observe four possible factor inter-correlations: Agreeableness-Openness, Extraversion-Conscientiousness, Extraversion-Neuroticism, and Conscientiousness-Neuroticism. We assessed structural invariance by examining changes in fit due to constraining these correlations to be equal across groups, as well as the magnitude of group estimates of each available correlation.

### Big Five and SES Indicators (H2)

We tested for positive associations of SES with conscientiousness, extraversion, and agreeableness, and negative association with neuroticism, through a series of bivariate correlations. We calculated the correlations between the four measured indicators of SES and the Big Five *observed scales* and reduced observed scales, resulting in twenty-eight bivariate correlations of interest. We conducted this analysis using the *psych* package in R (version  $\geq$  1.8.12; Revelle, 2018). In addition, we regressed each Big Five scale on education, self-reported income, and occupational prestige to assess the distinctive contribution of each SES variable. Because occupational income and occupational prestige are both based on the same item and were collinear in the exploratory data, we did not use occupational income and occupational prestige together in the same regression analyses.

### Age Trends by Socioeconomic Strata (H3)

The H3 analyses was conducted using linear regression models. We analyzed one Big Five trait and one SES indicator at a time, resulting in 20 different models. Each model consisted of a single Big Five trait regressed on age (mean-centered), an SES indicator (mean-centered), and an age-by-SES interaction term. These models estimated the

main effects of age and SES on personality, and an interaction effect between age and SES on personality. We report the main effects of age on personality and later discuss if the observed age trends, controlling for SES, are consistent with previous findings. The main outcome of interest in this analysis was the coefficient of the interaction term. A significant effect for the interaction term is interpreted as indicating that age trends in development for that trait vary as a function of SES.

### Missing Data Approach

The planned missingness design of the AIID study resulted in no participants responding to all Big Five measures. Instead, participants were randomly assigned to one of 20 individual difference measures. For participants assigned to Big Five measures, some provided responses to two trait measures, agreeableness and openness, and others to three trait measures, extraversion, conscientiousness, and neuroticism. We calculated observed scales scores across completed items and included participants who made one or more responses to trait items for any subsequent analysis which includes that trait, unless they were excluded for other reasons. In addition, to address missing data in the measurement invariance analysis (H1) we estimated the models using full information maximum likelihood (fiml).

For the correlation analysis (H2), we used pairwise deletion. Although this approach is generally not recommended it is acceptable if, as in this case due to the planned missingness design, the data is missing at random (MAR; Rubin, 1976). For the age trends analysis, we excluded participants who did not provide a response to the dependent variable of interest for each regression model. This resulted in participants being excluded from one analysis and included in another. For example, if an individual provided information about their income but not their education, they were included in the age trends regression model predicting income but not the model predicting education.

### Sample Size Considerations

The data for this study were collected prior to the design of the present study. Due to the planned missingness design, the subsets of the data that will be used for this analysis are considerably smaller than the total AIID sample size. We used all available data (after exclusions discussed earlier) for each analysis, resulting in different sample sizes for each.

In addition to the exploratory dataset, we had access to a masked version of the confirmatory dataset that replaced all nonmissing responses with the number 1. Because of uncertainty regarding whether the data underlying the 1s in the masked data would meet exclusion criteria, we could not a priori calculate an exact sample size for each analysis based on currently available information. However, estimating the sample sizes in the masked confirmatory data for each analysis suggests that they will be sufficient for valid inferences. We detail these considerations below:<sup>1</sup>

**H1 analyses.** For the measurement invariance analyses, we used the masked confirmatory data (all responses replaced in the data with a 1) to estimate the final sample

size for each analysis. Across all the SES indicators for each Big five trait, the estimated number of participants in each group ranged from  $N = 227$  to  $N = 1681$ . Sample-size planning in structural equation modeling (SEM), like that used in the measurement invariance analysis, is notably complex. However, for this analysis, the fit measure we used to assess invariance, MFI, is robust to sample sizes larger than 100 (Kang et al., 2016) and the  $d_{MACS}$  effect size, the metric we used to construct invariant scales, is relatively stable in smaller samples ( $>250$ ; Nye et al., 2019), supporting that we had enough data to draw valid inferences.

**H2 analyses.** For the correlation and regression analysis for H2, we anticipated there would be from  $N = 1763$  to  $N = 2192$  participants per pairwise observations for each of our 20 planned bivariate correlations, and between  $N = 1692$  and  $N = 1714$  participants for each of the 5 regression analyses. Correlations stabilize at  $\sim 250$  observations (Schönbrodt & Perugini, 2013) and thus we anticipated that this analysis afforded adequate precision. We also conducted a sensitivity analysis in G\*Power (Faul et al., 2009) to determine the smallest effect the planned analyses could detect. The results indicated that a fixed linear multiple regression model with three predictors (corresponding to 3 SES predictors), a sample size of 1692, desired power of 95%, and an alpha level of .05 could detect a small effect ( $f^2 = .01$ , or approximately  $R = .10$ ). Based on this sensitivity analysis, we concluded that the proposed analysis were adequately sensitive to detect effects at the smaller end of the range of those previously reported in the literature.

**H3 analyses.** For the age trends in personality moderated by SES analysis, we anticipated there would be from  $N = 1758$  to  $N = 2191$  participants for each linear regression model. A sensitivity analysis for a fixed linear multiple regression model, conducted in G\*Power (Faul et al., 2009), indicated that a model with a sample size of 2096, desired power of 95%, and an alpha of .05 could detect a small interaction effect ( $f^2 = 0.007$ ; equivalent to a  $\Delta R^2 = .007$  when adding the interaction term to the model). This demonstrates that the planned analyses were sensitive to detect interaction effects of meaningful size.

## Results

### Participants

After exclusions, the subset of data available for the confirmatory measurement invariance analysis (H1) included 6937 participants (age:  $M = 33$ ,  $SD = 13$ ; 67% female; 76% white) who provided responses to either the agreeableness and openness subscales ( $N = 3499$ ; age:  $M = 33$ ,  $SD = 13$ ; 67% female; 76% white) or the extraversion, conscientiousness, and neuroticism subscales of the BFI ( $N = 3438$ ; age:  $M = 33$ ,  $SD = 13$ ; 67% female; 75% white). After excluding participants who indicated they were students and those under

25 years of age for the remainder of the analyses (H2 and H3), 4388 participants (age:  $M = 39$ ,  $SD = 11$ ; 65% female; 77% white) remained in the two groups, with 2220 who responded to the agreeableness and openness subscales (age:  $M = 39$ ,  $SD = 11$ ; 65% female; 78% white) and 2168 who responded to the extraversion, conscientiousness, and neuroticism subscales (age:  $M = 40$ ,  $SD = 11$ ; 66% female; 77% white). As discussed above, sample sizes for each individual analysis may differ slightly depending on missingness for the variables included in the model.

Descriptive results for the Big Five traits from the AIID confirmatory dataset (students and under 25 excluded) are reported in Table 1 and for SES indicators in Table 2. For the Big Five traits, the means and standard deviations do not suggest floor or ceiling effects and internal consistency coefficients are in the conventionally acceptable range, similar to previous studies using the BFI (John & Srivastava, 1999).

### Measurement Invariance (H1)

Using the  $\Delta MFI > .01$  criterion as an indication of measurement noninvariance, results demonstrated that extraversion was noninvariant across different levels of self-reported income, occupational income, and occupational prestige (see Table 3). Agreeableness was invariant across all SES indicators. Conscientiousness was noninvariant across education and occupational prestige and invariant across income and occupational income. Neuroticism was noninvariant across levels of education, self-reported income, and occupational income; it was invariant across occupational prestige. Openness was noninvariant across all SES indicators. We conducted these analyses using both the default reference indicators in *lavaan*, as well as the reference indicators selected by the procedure discussed in the Analysis Plan section. These analyses resulted in nearly identical outcomes.

To specify partially invariant factors, we freed model constraints (e.g., equality of intercepts or residual variance) based on the Lagrange multiplier scores produced by *lavaan* and  $d_{MACS}$  coefficients until the model comparisons satisfied the  $\Delta MFI$  criterion. In most cases, we were able to create partially invariant factors by removing constraints on one or two items, with two exceptions (i.e., neuroticism across self-reported income, and openness across education categories) which required freeing constraints on four and three items, respectively. See Table 4 for the list of items and freed constraints. We then created invariant scale scores by excluding the items that were contributing to noninvariance. We refer to these as the *reduced observed* scale scores, and the originally published scales as *observed* scale scores.

**Structural Invariance.** In addition to assessing mea-

1 In the stage 1 Registered Report, we included 6 tables detailing the expected sample size for each of our planned analyses. In this stage 2 Registered Report, we moved these tables to the Supplement (see Supplemental Tables S1 to S6 for details) and adapted each to indicate the predicted and observed samples sizes. Some of the predicted sample sizes were smaller than the observed sample sizes due to a coding error in the analysis predicting the sample sizes from the masked data. The purpose of reporting the predicted sample sizes was to show the planned analyses offered adequate sensitivity to test the preregistered hypotheses. Because this error resulted in a greater number of participants being included in the confirmatory analyses all analyses conducted were adequately sensitive.

**Table 2. Correlations Between Socioeconomic Indicators**

Variable	M	SD	1	2	3
Income	3.15	1.16			
Education	3.04	0.87	.28*		
			[.26, .31]		
Occupational Prestige	55.65	9.69	.30*	.45*	
			[.27, .32]	[.42, .47]	
Occupational Income	55.05	21.59	.36*	.35*	.75*
			[.34, .39]	[.32, .38]	[.74, .77]

Note: The correlations were calculated from data with students and participants under 25 years of age excluded ( $N = 4388$ ) using pairwise deletion, sample sizes for each vary. Income and education variables were reported on a 1-5 scale and occupational prestige and income were measured on a 0-100 scale, for both higher numbers equals higher SES. M and SD represent mean and standard deviation, respectively and values in square brackets indicate the 95% confidence interval for each correlation. \* indicates  $p < .001$ .

surement invariance, we also assessed the structural invariance between SES groups. Due to the nature of the data, we were only able to examine structural invariance in subsets of traits: one set of models investigated the covariance between agreeableness and openness, and another set examined the covariances among extraversion, conscientiousness, and neuroticism. To estimate structural invariance, we fit two multigroup CFA models for each SES indicator in the two groups of traits. In the invariant model we constrained the covariances between latent traits to be equal across SES groups, and in the non-invariant model we allowed these covariances to be estimated freely. A comparison of AIC between the invariant and non-invariant models showed that 5 of the 8 model comparisons favored the invariant model and three were equivocal ( $\Delta AIC < 1$ ).

### Big Five and SES Indicators (H2)

We analyzed the associations between personality traits and the four SES indicators in two ways. First, we examined zero-order correlations, reported in Table 5. Second, to assess the distinctive contribution of each SES indicator we estimated multiple linear regression models, in which each Big Five scale was regressed on education, self-reported income, and occupational prestige in a single model; these are reported in Table 6. Support for the trait-specific hypotheses was mixed and inconsistent across the 4 SES indicators. Conscientiousness was hypothesized to have a positive association with SES; in this data it had a significant, small correlation with income ( $r = .10$ ) and a suggestive one with educational attainment, but it showed no association with occupational income or prestige. Extraversion was also hypothesized to have a positive association with SES; it had significant but small correlations with income ( $r = .12$ ) and occupational income ( $r = .08$ ), a suggestive one with educational attainment, but it showed no association with occupational prestige. Neuroticism was hypothesized to have a negative association with SES; it had a significant correlation in the predicted direction with income ( $r = -.16$ ) and occupational income ( $r = -.10$ ) and a suggestive one with both educational attainment and occupational prestige. There was no support for the hypothesis that agreeableness would be positively associated with SES, but there was a signifi-

cant negative correlation between agreeableness and occupational income ( $r = -.08$ ) and a suggestive negative correlation with occupational prestige. These associations were similar to the standardized regression coefficients, suggesting relatively little overlap among the SES indicators in predicting personality traits.

Because the measurement invariance analyses indicated noninvariance across some SES categories, we should use caution when interpreting the observed scale score zero-order correlations and regression betas. To address this issue, we ran the correlation and regression analyses with the reduced observed scales (reported in Tables 5 and 6). As an additional way of looking at this question, we ran non-pre-registered analyses with partially invariant latent models using a CFA approach. In those models, we compared latent factor means for each level of the SES categories. Across these three approaches, the associations between Big Five traits and SES were similar. The analyses with reduced observed scales were similar to the a priori observed scales, and latent factor mean comparisons led to the same qualitative conclusions for virtually all of the Big Five domains and SES categories (see Supplemental Tables S7 to S14 for details). Together, these results largely suggest that the observed correlations between SES and Big Five traits are not driven by measurement artifacts.

### Age Trends by SES (H3)

We ran 20 regression models, each predicting a Big Five trait from a main effect of age, a main effect of SES (with the 4 indicators in separate models), and an age-by-SES interaction. Parameter estimates for the age-trend linear regression models are presented in Table 7.

**Main effects of age.** The results for the main effects of age on Big five traits were relatively consistent when comparing models with the different indicators of SES. Extraversion was the only trait that had a minor inconsistency in the statistical significance of coefficients, with the model controlling for occupational income showing an effect that just crosses the threshold for suggestive ( $b_{age} = 0.87$ ,  $p = .049$ ) and the others nonsignificant ( $range_b: 0.20$  to  $0.81$ ). There was a consistent significant but small positive effect of age on Agreeableness ( $range_b: 1.65$  to  $2.13$ ,  $p < .005$ ). This

**Table 3. Results of the Measurement Invariance Analyses of Big 5 across Different SES Indicators**

Personality Domain	SES Indicator	Model	df	MFI	RMSEA	AIC
Extraversion	Education	configural	100	.839	.133	80360.44
		metric	128	.835	.119	80363.72
		scalar	156	.830	.109	80375.06
		full	188	.821	.103	80411.97
	Self-reported Income	configural	120	.839	.132	81180.03
		metric	155	.838	.117	81158.27
		scalar	190	.824*	.111	81233.13
		full	230	.817	.103	81242.18
	Occupational Income	configural	80	.837	.134	80441.71
		metric	101	.835	.120*	80434.04
		scalar	122	.825*	.112	80488.37
		full	146	.821	.104	80500.80
	Occupational Prestige	configural	80	.836	.134	81193.95
		metric	101	.834	.120	81194.09
		scalar	122	.823*	.113	81256.80
		full	146	.816	.105	81284.24
Agreeableness	Education	configural	135	.886	.095	90998.17
		metric	167	.883	.086	90992.58
		scalar	199	.877	.081	91000.22
		full	235	.869	.077	91027.02
	Self-reported Income	configural	162	.875	.100	90080.78
		metric	202	.873	.090	90049.86
		scalar	242	.870	.083	90034.35
		full	287	.863	.078	90041.98
	Occupational Income	configural	108	.880	.097	90459.11
		metric	132	.879	.088	90442.10
		scalar	156	.877	.082	90432.53
		full	183	.873	.077	90440.71
	Occupational Prestige	configural	108	.885	.095	91706.58
		metric	132	.885	.086	91681.15
		scalar	156	.882	.080	91682.21
		full	183	.877	.076	91688.53
Conscientiousness	Education	configural	135	.877	.099	87042.8
		metric	167	.876	.089	87020.31

Personality Domain	SES Indicator	Model	df	MFI	RMSEA	AIC	
Neuroticism		scalar	199	.859*	.087	87109.8	
		full	235	.848*	.084	87152.4	
	Self-reported Income	configural	162	.886	.095	88026.51	
		metric	202	.884	.086	88001.44	
		scalar	242	.877	.081	88014.6	
		full	287	.868	.077	88033.69	
	Occupational Income	configural	108	.885	.095	87185.27	
		metric	132	.883	.087	87171.99	
		scalar	156	.876	.083	87204.67	
		full	183	.870	.078	87215.01	
	Occupational Prestige	configural	108	.882	.096	88009.52	
		metric	132	.882	.087	87990.02	
		scalar	156	.872*	.084	88037.10	
		full	183	.867	.079	88046.15	
	Openness	Education	configural	100	.878	.114	82528.14
			metric	128	.877	.101	82507.76
			scalar	156	.866*	.096	82561.7
			full	188	.861	.089	82568.32
		Self-reported Income	configural	120	.880	.113	83330.29
			metric	155	.879	.100	83302.56
scalar			190	.862*	.097	83387.84	
full			230	.860	.089	83367.64	
Occupational Income		configural	80	.876	.115	82693.19	
		metric	101	.876	.103*	82677.27	
		scalar	122	.866*	.097	82725.84	
		full	146	.865	.089	82709.14	
Occupational Prestige		configural	80	.881	.113	83577.43	
		metric	101	.880	.101*	83565.18	
		scalar	122	.872	.095	83603.10	
		full	146	.872	.087	83577.60	
Openness		Education	configural	175	.785	.118	94518.28
			metric	211	.782	.108	94507.5
			scalar	247	.754*	.107	94706.96
			full	287	.738*	.103	94807.71
	Self-reported Income	configural	210	.785	.118	93746.62	
		metric	255	.782	.107	93721.58	

Personality Domain	SES Indicator	Model	df	MFI	RMSEA	AIC
		scalar	300	.759*	.105	93874.55
		full	350	.755	.098	93859.06
	Occupational Income					
		configural	140	.784	.118	93915.37
		metric	167	.783	.108	93899.92
		scalar	194	.759*	.107	94068.85
		full	224	.754	.101	94086.48
	Occupational Prestige					
		configural	140	.785	.117	95254.39
		metric	167	.783	.108	95244.11
		scalar	194	.763*	.106	95387.00
		full	224	.759	.099	95397.83

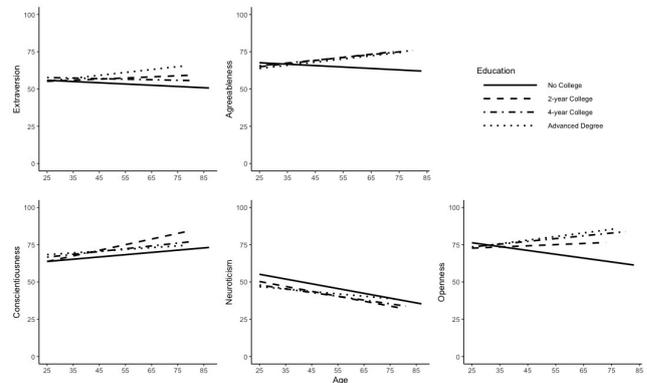
\* ΔMFI > .01

effect size is equivalent to roughly a 2 POMP unit increase in agreeableness for every decade lived. There were also significant positive main effects of age on both conscientiousness (range<sub>b</sub>: 1.88 to 2.46,  $p < .005$ ) and openness (range<sub>b</sub>: 1.05 to 1.81,  $p < .005$ ) with similar effect sizes, suggesting about a 2 POMP unit increase per decade. Age had the opposite effect on neuroticism with consistent, significant, negative effects across the models (range<sub>b</sub>: -2.67 to -1.79,  $p < .005$ ), suggesting a 2 POMP decrease per decade.

**Main effects of SES.** The estimates of the main effects of the SES indicators included in these models are consistent with the zero-order correlations and multiple regression models presented in Tables 5 and 6 above. The only notable departure from these results was for the models with educational attainment.

In the H2 models there was only a significant association between educational attainment and openness, but in the H3 models that control for age there was not a significant association with openness ( $b_{edu} = 0.39, p < .005$ ) and there was a suggestive positive association with conscientiousness ( $b_{edu} = 1.59, p = .005$ ) and a suggestive negative association with neuroticism ( $b_{edu} = -1.67, p = .046$ ).

**Age by SES interactions.** The main focus of the H3 analyses was on whether age effects on personality traits would be moderated by SES. Overall there was not much support for age-by-SES interactions, with only one suggestive and one significant interaction effects across the 20 models. The three specific hypotheses based on normative role-timing received little support. There was no support that SES is positively associated with the age slope for neuroticism (H3a). For conscientiousness (H3b), only the model with educational attainment supported the hypothesis that SES is negatively associated with the age slope, and only at the suggestive level ( $b_{ageXedu} = -0.86, p = .016$ ). Figure 2 shows that, descriptively, this effect may reflect a large increase in conscientiousness by age for those with 2-year degrees rather than a monotonic interaction. For openness (H3c), there was one significant interaction of age and SES, but the effect was in the opposite direction than was hypothesized. It was the interaction between age and educa-



**Figure 2. Age Trends in Personality Development as a Function of Educational Attainment**

Note: BFI traits are displayed in POMP units. Figure created with ggplot2 package (Wickham, 2016)

tional attainment ( $b_{ageXed\_att} = 1.07, p < .005$ ), such that openness was more strongly associated with age with higher educational attainment. Figure 2 shows that there was a negative association between openness and age for people with no college but a positive association between openness and age in the other groups.

Parameter estimates from the same analyses conducted with reduced observed scales were mostly consistent (reported in Supplemental Table S15). There was one exception for the age by SES interactions; the age by education effect on neuroticism was not significant in the observed scale models ( $b_{ageXedu} = 0.74, p = .073$ ) but suggestive in the reduced observed scale models ( $b_{ageXedu} = 0.87, p = .044$ ).

**Table 4. Items That Were Removed to Create Partial Invariance Scores**

Trait	Noninvariance Category	Item	Freed Constraint
Extraversion	Self-reported Income	"Is full of energy."	Free intercept
		<b>"Is sometimes shy, inhibited."</b>	Free intercept
	Occupational Income	"Is talkative"	Free intercept
	Occupational Prestige	<b>"Is sometimes shy, inhibited."</b>	Free intercept
Conscientiousness	Education	"Is a reliable worker."	Free intercept
		"Does things efficiently."	Free intercept
	Occupational Prestige	"Can be somewhat careless"	Free intercept
Neuroticism	Education	<b>"Remains calm in tense situations."</b>	Free intercept
		"Gets nervous easily."	Free intercept
	Self-reported Income	<b>"Is depressed, blue."</b>	Free intercept
		"Is emotionally stable, not easily upset."	Free intercept
		"Can be moody."	Free intercept
	Occupational Income	<b>"Remains calm in tense situations."</b>	Free intercept
Occupational Income	<b>"Is depressed, blue."</b>	Free Intercept	
Openness	Education	<b>"Values artistic, aesthetic experiences."</b>	Free residual variance
		<b>"Prefers work that is routine."</b>	Free intercept and residual variance
		"Likes to reflect, play with ideas."	Free residual variance
	Self-reported Income	"Is inventive."	Free intercept
		<b>"Prefers work that is routine."</b>	Free intercept
	Occupational Prestige	<b>"Values artistic, aesthetic experiences."</b>	Free intercept
		<b>"Prefers work that is routine."</b>	Free intercept
	Occupational Income	<b>"Values artistic, aesthetic experiences."</b>	Free residual variance
Occupational Income	<b>"Prefers work that is routine."</b>	Free intercept	

Note: Items in bold contributed to noninvariance across more than one SES category.

**Table 5. Zero-order Correlations Between Observed Scale Scores and Reduced Observed Scale Scores and SES Indicators**

Trait	Income	Education	Occupational Prestige	Occupational Income
Extraversion	.12*** (.10***)	.05*	.02 (.02)	.08*** (.09***)
Agreeableness	.02	-.00	-.06*	-.08***
Conscientiousness	.10***	.04* (.06*)	-.01	.03 (.02)
Neuroticism	-.16*** (-.11***)	-.05* (-.04)	-.06*	-.10*** (-.09***)
Openness	.02 (-.02)	.11*** (.07***)	.04 (.02)	.03 (.01)

Note: Correlations in parentheses indicate the correlations with reduced observed scale scores and SES indicators. \*p < .05, \*\*\*p < .005

**Table 6. Standardized Regression Coefficients from Models Regressing Observed and Reduced Observed Scores of Big Five Traits on three indicators of SES**

Trait	SES Predictors		
	Income	Education	Occupational Prestige
Extraversion	.11***	.03	-.02
<i>Income</i>	.09***	.03	-.03
<i>Occ Prestige</i>	.11***	.04	-.03
<i>Occ Income</i>	.12***	.02	-.02
Agreeableness	.04	.04	-.08***
Conscientiousness	.10***	.03	-.05
<i>Education</i>	.11***	.05	-.04
<i>Occ Income</i>	.10***	.04	-.05*
Neuroticism	-.14***	.03	-.03
<i>Education</i>	-.14***	.04	-.03
<i>Income</i>	-.09***	.02	-.02
<i>Occ Income</i>	.12***	.03	-.03
Openness	-.00	.10***	-.00
<i>Education</i>	-.01	.07***	-.00
<i>Income</i>	-.04	.09***	-.02
<i>Occ Prestige</i>	-.01	.08***	-.01
<i>Occ Income</i>	-.01	.08***	-.01

Note: Models in italics are with reduced observed scores for that trait and indicator of SES as the dependent variable. \*p < .05, \*\*\*p < .005

**Table 7. Unstandardized Regression Coefficients from Age Trend Linear Regression Models (H3)**

	Big Five Personality Trait (Dependent Variable)				
	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness
<i>Self-report Income</i>					
Age	0.20 (0.44)	1.83*** (0.33)	1.88*** (0.35)	-1.79*** (0.40)	1.67*** (0.28)
Income	1.79*** (0.52)	-0.36 (0.39)	1.19** (0.42)	-2.06*** (0.48)	-0.44 (0.34)
Age X Income	0.27 (0.35)	0.26 (0.26)	-0.16 (0.28)	-0.10 (0.33)	0.27 (0.23)
Observations	2,081	2,106	2,081	2,081	2,106
<i>Educational Attainment</i>					
Age	0.39 (0.40)	1.65*** (0.35)	2.46*** (0.36)	-2.67*** (0.41)	1.05*** (0.30)
Education	0.36 (0.71)	-0.66 (0.52)	1.59*** (0.57)	-1.67* (0.65)	0.39 (0.45)
Age X Education	0.72 (0.44)	0.35 (0.34)	-0.86* (0.36)	0.74 (0.41)	1.07*** (0.29)
Observations	2,162	2,213	2,162	2,162	2,213
<i>Occupational Prestige</i>					
Age	0.81 (0.44)	2.10*** (0.33)	2.44*** (0.35)	-2.55*** (0.0)	1.81*** (0.29)
Occ Prestige	-0.01 (0.06)	-0.12* (0.04)	0.00 (0.05)	-0.06 (0.05)	0.03 (0.04)
Age X Occ Prestige	0.07 (0.05)	0.03 (0.03)	-0.03 (0.04)	-0.06 (0.04)	0.03 (0.03)
Observations	1,999	2,056	1,999	1,999	2,056
<i>Occupational Income</i>					
Age	0.87* (0.44)	2.13*** (0.32)	2.43*** (0.35)	-2.58*** (0.40)	1.77*** (0.28)
Occ Income	0.07* (0.03)	-0.08*** (0.02)	0.03 (0.02)	-0.06* (0.02)	0.00 (0.02)
Age X Occ Income	-0.00 (0.02)	0.02 (0.01)	-0.02 (0.02)	-0.03 (0.02)	0.02 (0.01)
Observations	1,981	2,027	1,981	1,981	2,027

Note: Standard error in parentheses. \*p < .05, \*\*\*p < .005

## Discussion

Can the Big Five be validly compared across SES? If so, what do those comparisons reveal about associations and age trends? First, using a large archival dataset, we tested the measurement invariance of a popular Big Five instrument, the BFI, across levels of several SES indicators. Second, we estimated the associations between personality traits and SES indicators. Third, we tested whether age trends in personality were moderated by SES. The results did not radically overturn any conclusions from previous studies, but they do paint a complex picture of how the Big Five intersect with SES. Here we discuss implications for

measuring the Big Five in socioeconomically diverse samples, for research programs that assume meaningful associations between the Big Five and SES, and for theories of how personality change intersects with social and economic conditions.

## Measurement Invariance and Its Practical Consequences

Across 20 measurement invariance analyses, 12 showed evidence of non-invariance using the  $\Delta\text{MFI} > .01$  threshold. Thus, a first pass at this question would suggest that people with different socioeconomic backgrounds may not be an-

swering the BFI in ways that allow for valid comparisons. However, the authors who suggested this threshold also cautioned against taking it too seriously (Kang et al., 2016). Thus, rather than stopping at this result, it prompted us to take a closer look at the data in two ways: by examining the items that were responsible for noninvariance, and by gauging the practical impact of correcting for it.

**Examining noninvariant items.** In examining the non-invariant items (see Table 4), one observation we made was that two items with “work” or “worker” were noninvariant. One possibility is that participants differ in the extent to which they are thinking of paid labor, versus other meanings such as schoolwork or anything involving effort. Furthermore, research on the meaning of work (in the sense of paid labor) suggests that financial circumstances, and the nature of work tasks, change the way people view work and connect it to their sense of self (Rosso et al., 2010). So even among participants thinking about employment, systematic differences in the meaning of work could have been responsible for these differences.

A second observation was that “values artistic, aesthetic experiences” was noninvariant. Duarte (2015) suggested that several BFI Openness items, including that one, reflect “intellectualism and urban sophistication” and may not validly reflect openness in all populations. Duarte proposed an urban-rural divide as a key distinction in how people would respond to such items, but his characterization of “rural” in this context emphasized the rural working class, so an SES-linked difference in how participants respond to this item would be aligned with his critique.

A third observation was that the neuroticism scale fared poorly across self-reported income, with 4 items showing noninvariance. Events that are associated with lower income, such as a history of episodes of unemployment, have been associated with reductions in well-being, which is closely associated with neuroticism (Lucas et al., 2016). Different items reflect different facets of neuroticism, and if the effects of unemployment are unevenly distributed across facets, that could have led to a pattern of noninvariance.

The present study examined measurement invariance for the original BFI (John & Srivastava, 1999), which has been revised to the BFI-2 (Soto & John, 2017). The revision included modifying, removing, and replacing items. Across the measurement invariance analysis, we identified 15 items that were noninvariant across one of the indicators of SES. From the BFI to the BFI-2, five of these items remain unchanged, 5 were modified, and 5 were removed. We therefore cannot draw firm conclusions about whether the BFI-2 is invariant across SES; it will be important to investigate the revised scale in future studies.

**Practical impact of correcting for noninvariance.** In places where the measurement analyses indicated noninvariance, we corrected for it by creating reduced observed scales and partially invariant factors, and then we evaluated

whether the associations between the Big Five and SES looked different after these corrections. These analyses largely showed that noninvariance did not make much difference for substantive conclusions: in nearly every case, a researcher would have reached similar conclusions whether they used a priori scales or corrected ones. We suggest that this can increase confidence that past findings do not result from measurement artifacts, but we do so with caveats. Noninvariance might matter more in other settings or samples, or for more focused hypotheses about the traits that showed it. Thus, rather than sounding the all-clear, we offer three recommendations. First, future researchers should routinely test for invariance when studying personality traits and SES, or when studying personality with other constructs in socioeconomically diverse samples. For the common practice of analyzing correlations with observed scale scores, this will require testing for full (strict) invariance. Second, given the high sensitivity of conventional thresholds to even trivial noninvariance, we also recommend that such tests should include an assessment of practical impact, for example by calculating effect sizes or comparing results with and without corrections. Third, given there were multiple noninvariant items across domains, we recommend the use of longer form versions of personality measures, when possible, over short form or abbreviated scales. The inclusion of additional items will provide researchers a safety margin to conduct measurement invariance analyses in new datasets and remove some noninvariant items if necessary.

### Associations Between SES and the Big Five

In previous studies, SES indicators have been positively associated with extraversion, agreeableness, and conscientiousness, and negatively associated with neuroticism, with absolute correlations ranging from .10 to .30. In the present analyses, however, only 4 correlations (out of 16 predicted) were .10 or larger, and the strongest correlation was only -.16. Directionally, the signs of the correlations were mostly consistent with predictions, though the correlations of two occupation-derived indices with agreeableness were in the opposite direction.<sup>2</sup> Overall, then, the pattern of results was weaker and somewhat less consistent than a reading of at least some earlier studies would lead one to expect.

What could explain the difference from previous results? One possibility to consider first is data quality. Random noise attenuates correlations, so random responding or other issues with the AIID dataset could have led to the lower correlations. However, we would expect random noise to show up in other ways, such as lower-than-usual internal consistency coefficients for Big Five scales, which we did not observe here. Therefore, we do not think the weaker correlations reflect a serious data-quality issue. A related consideration is the coarseness of the measurement scales, which can attenuate effect sizes (Aguinis et al., 2009). Par-

<sup>2</sup> Goldberg et al. (1998) found a negative correlation between educational attainment and agreeableness. This study was brought to our attention after conducting analyses, and therefore was not reflected in our preregistered hypotheses.

participants chose from a set of relatively broad occupation categories rather than specific occupations. Income, a continuous variable, was measured by asking participants to select from binned ranges. This is a common practice and therefore may not explain differences with previous studies that used a similar approach, but future studies may benefit from more finer-grained measurement of these variables.

Another possibility is that effect magnitudes in previous studies may have been upwardly biased by analytic flexibility or publication bias (Open Science Collaboration, 2015; Zwaan et al., 2018). The current study included safeguards against those sources of bias through the use of preregistration and the registered report publication process. An additional piece of evidence for this conclusion comes from a study of a large, nationally representative that was published after the Stage 1 manuscript was accepted. Zisman & Ganzach (2020) reported associations between the Big Five and SES indicators but did not make them the focus of hypotheses, making those results less likely to have been inflated by publication bias. In that study, correlations between a Big Five measure and two SES measures used here, educational attainment and income, were similar in magnitude and direction to the present results. Taking those results and the present ones together, it appears that the associations between the Big Five and SES are quite small – and perhaps smaller than what researchers have previously believed.

In light of this pattern of weak associations, it may be important to revisit previous calls for studying personality traits in order to understand, and potentially intervene on, educational and economic outcomes (Bleidorn et al., 2019; Duckworth & Gross, 2014; Heckman et al., 2006). It would be premature to abandon such efforts. But it may be valuable to reexamine some proposals, such as the idea that interventions to increase traits like conscientiousness, or highly overlapping constructs like grit (Credé et al., 2017), will have downstream effects on outcomes like income or employment. Such proposals make sense if effect sizes are pragmatically large enough to make the intervention worth doing. The present results are not a conclusive refutation, but they do raise reasons for caution and further investigation before proceeding with interventions.

## Age Trends and SES

Most of the main effects of age on personality traits were consistent with previous studies. Agreeableness, conscientiousness, and openness increased with age; neuroticism decreased; and extraversion was relatively flat. Of these findings, only the age trend for openness was at odds with previous cross-sectional and longitudinal studies, in which openness is typically either flat or decreasing (Roberts et al., 2006; Srivastava et al., 2003). We do not have a clear explanation for this pattern, but if the observed interaction between age and education is a true effect then it might be explained by differences in education levels of the samples. Another possibility could be sampling bias, specifically an openness-by-selection effect. The data were from a website for people interested in taking the IAT; perhaps as age increases, higher openness becomes a stronger determinant of interest in taking online psychological tests. However, we

cannot say with certainty whether that explains the pattern.

A major question for this investigation was whether age trends vary by SES. In general, the answer was no: age trends were mostly parallel across different levels of SES, as indicated by the general lack of substantial interactions. In this reasonably well-powered design, significance tests showed only two suggestive or significant interactions out of 20 tests: differences in the age-openness slope across education, and differences in the age-conscientiousness slope across education. Of these, the age-by-education interaction for openness was the only one that reached significance at the .005 level. The pattern could be interpreted through a developmental lens as suggesting that education promotes growth in openness. However, education is not static across the lifespan, and higher openness may be associated with seeking out educational opportunities. Thus, the age trend could also be a product of cumulative selection effects rather than maturation.

Three of the age trend analyses were tests of specific hypotheses that were derived from previous research on social investment (Bleidorn et al., 2013). The present results were not strongly supportive of these hypotheses. The predicted association between SES and age slopes was null for neuroticism. For conscientiousness, the data supported the hypothesis only for education, not the other SES indicators. For openness, three indicators were null and one indicator was significant in the opposite of the hypothesized direction. These results are far from a definitive falsification of social investment theory writ large, which should be evaluated in light of an overall body of findings rather than a single study. However, they do raise questions about whether there are robust role-timing effects on personality that can be measured and analyzed this way.

Overall these results should give some reassurance that claims about mean-level or “normative” personality development are generalizable across SES. However, this conclusion comes with notable limitations. Age trends in a cross-sectional design like this one can reflect maturation, cohort effects, or age-dependent sampling bias, and we cannot definitively say which is responsible. Interpretations of the interaction results are further complicated by the fact that SES is not stable across the lifespan: people can change occupations, change income, and increase in educational attainment with age. Thus, we interpret the present results to offer modest, but not definitive, evidence of the generalizability of age trends.

## Limitations and Conclusions

This study had several limitations. Measurement invariance analyses are a useful tool, but they have been criticized as too sensitive, potentially detecting minor violations of invariance that can be safely ignored (Funder, 2020). The analyses are conditional on a factor model that may be useful but is almost certainly incorrect (Srivastava, 2020; Wood et al., 2015). We attempted to address the former issue by gauging the practical differences between a priori and corrected measures, but the psychometric properties of Big Five measures should never be considered settled for all populations and applications.

The analyses of this study focused on a single instru-

ment, the BFI, and cannot speak directly to other instruments. The BFI consists of short phrases based on adjectives identified in lexical analyses, and as a result the items are moderately abstract (John & Srivastava, 1999). Other instruments, such as the NEO PI-R (Costa & McCrae, 1992) or the IPIP-NEO (Goldberg, 1999), include more concrete statements. Items like “[I] avoid philosophical discussions” or “do just enough work to get by” (both from the IPIP-NEO) may mean different things to people with different educational experiences or in different workplace contexts. Whether those and other instruments have problematic levels of noninvariance across SES indicators remains to be seen.

The AIID sample, though large, is a convenience sample and therefore presents some constraints on the generality of the results. Participants enrolled in the study based on their interest in taking an Implicit Association Test. Data-quality checks and consistencies with previous studies gave some reassurance, but we cannot rule out that some results may be artifacts of self-selection bias. Because participants self-selected into the study, the sample is not demographically representative of the United States. It is more white, female, and better educated than the general population, which may limit the generalizability of the findings. In addition, the AIID study’s cross-sectional design means that age trends cannot be interpreted unambiguously as reflecting mean-level change, as opposed to cohort differences or age-dependent sampling bias.

Despite these limitations, we believe this study should give researchers some modest reassurance that it is possible to make valid comparisons of Big Five traits across levels of SES. The small effect sizes suggest that the story is largely one of similarity. That might temper some expectations about the potential for personality interventions to change economic outcomes. It also suggests boundaries on how much we can attribute social inequality to differences in Big

Five personality traits, or conversely, to how much social inequality leads to changes in personality.

More broadly, these results should give personality researchers more confidence to study socioeconomically diverse samples. It is reasonable to wonder whether it is appropriate to use personality instruments that relied heavily on college-student samples in their development. The present results suggest some reasons for caution, but also that that is not a fatal flaw. Given the considerable advantages of more diverse and inclusive sampling, researchers who are attentive to good measurement can continue to work toward expanding the scope of personality research.

---

## Contributions

Contributed to conception and design: BTH, CC, JP, PR, CBP, RML, SS. Contributed to analysis and interpretation of data: BTH, CC, JP, PR, CBP, RML, SS. Drafted and/or revised the article: BTH, CC, JP, PR, RML, CBP, JP, SS. Approved the submitted version for publication: BTH, CC, JP, PR, RML, CBP, SS.

## Competing interests

The authors have no competing interests.

## Data accessibility statement

The AIID study data can be found here: <https://osf.io/gvzwm/>.

The Stage 1 Registered Report and R code for the analyses for this manuscript are publicly available on OSF: <https://osf.io/wv59a/>

Submitted: January 04, 2021 PDT, Accepted: May 25, 2021 PDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license’s legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## REFERENCES

- Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: the challenge of the gradient. *American Psychologist*, *49*(1), 15–24. <https://doi.org/10.1037/0003-066x.49.1.15>
- Aguinis, H., Pierce, C. A., & Culpepper, S. A. (2009). Scale coarseness as a methodological artifact: Correcting correlation coefficients attenuated from using coarse scales. *Organizational Research Methods*, *12*(4), 623–652. <https://doi.org/10.1177/1094428108318065>
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6–10. <https://doi.org/10.1038/s41562-017-0189-z>
- Bleidorn, W., Hill, P. L., Back, M. D., Denissen, J. J. A., Hennecke, M., Hopwood, C. J., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Wagner, J., Wrzus, C., Zimmermann, J., & Roberts, B. (2019). The policy relevance of personality traits. *American Psychologist*, *74*(9), 1056–1067. <https://doi.org/10.1037/amp0000503>
- Bleidorn, W., Klimstra, T. A., Denissen, J. J., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social-investment theory. *Psychological Science*, *24*(12), 2530–2540. <https://doi.org/10.1177/0956797613498396>
- Borghuis, J., Denissen, J. J. A., Oberski, D., Sijtsma, K., Meeus, W. H. J., Branje, S., Koot, H. M., & Bleidorn, W. (2017). Big Five personality stability, change, and codevelopment across adolescence and early adulthood. *Journal of Personality and Social Psychology*, *113*(4), 641–657. <https://doi.org/10.1037/pspp0000138>
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*(3), 425–440. <https://doi.org/10.1007/s11336-006-1447-6>
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371–399. <https://doi.org/10.1146/annurev.psych.53.100901.135233>
- Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, *37*(4), 319–338. [https://doi.org/10.1016/s0092-6566\(02\)00578-0](https://doi.org/10.1016/s0092-6566(02)00578-0)
- Chapman, B. P., Fiscella, K., Kawachi, I., & Duberstein, P. R. (2010). Personality, Socioeconomic Status, and All-Cause Mortality in the United States. *American Journal of Epidemiology*, *171*(1), 83–92. <https://doi.org/10.1093/aje/kwp323>
- Clark, D. A., Listro, C. J., Lo, S. L., Durbin, C. E., Donnellan, M. B., & Neppl, T. K. (2016). Measurement invariance and child temperament: An evaluation of sex and informant differences on the Child Behavior Questionnaire. *Psychological Assessment*, *28*(12), 1646–1662. <https://doi.org/10.1037/pas0000299>
- Conger, R. D., Wallace, L. E., Sun, Y., Simons, R. L., McLoyd, V. C., & Brody, G. H. (2002). Economic pressure in African American families: A replication and extension of the family stress model. *Developmental Psychology*, *38*(2), 179–193. <https://doi.org/10.1037/0012-1649.38.2.179>
- Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5–13. <https://doi.org/10.1037/1040-3590.4.1.5>
- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, *113*(3), 492–511. <https://doi.org/10.1037/pspp0000102>
- Damian, R. I., Su, R., Shanahan, M., Trautwein, U., & Roberts, B. W. (2015). Can Personality Traits and Intelligence Compensate for Background Disadvantage? Predicting Status Attainment in Adulthood. *Journal of Personality and Social Psychology*, *109*(3), 473–489. <https://doi.org/10.1037/pspp0000024>
- del Barrio, V., Carrasco, M. A., & Holgado, F. P. (2006). Factor structure invariance in the Children's Big Five Questionnaire. *European Journal of Psychological Assessment*, *22*(3), 158–167. <https://doi.org/10.1027/1015-5759.22.3.158>
- Diemer, M. A., Mistry, R. S., Wadsworth, M. E., López, I., & Reimers, F. (2013). Best Practices in Conceptualizing and Measuring Social Class in Psychological Research. *Analyses of Social Issues & Public Policy*, *13*(1), 77–113. <https://doi.org/10.1111/asap.12001>

- Duarte, J. L. (2015, February 10). *I was denied admission to a PhD program because of my perceived political views: Reflections of a sellout; how diversity would strengthen social science* [Blog post]. <http://www.joseduarte.com/1/post/2015/02/i-was-denied-admission-to-a-phd-program-because-of-my-perceived-political-views-reflections-of-a-sellout-how-diversity-would-strengthen-social-science.html>
- Duckworth, A., & Gross, J. J. (2014). Self-Control and Grit: Related but Separable Determinants of Success. *Current Directions in Psychological Science*, 23(5), 319–325. <https://doi.org/10.1177/0963721414541462>
- Dueber, D. (2019). *dmacs: Measurement Nonequivalence Effect Size Calculator*. <https://CRAN.R-project.org/package=dmacs>
- Epstein, W. M. (2006). Response bias in opinion polls and American social welfare. *The Social Science Journal*, 43(1), 99–110. <https://doi.org/10.1016/j.sosci.2005.12.010>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <https://doi.org/10.3758/brm.41.4.1149>
- Flake, J. K., Pek, J., & Hehman, E. (2017). Construct Validation in Social and Personality Research: Current Practice and Recommendations. *Social Psychological and Personality Science*, 8(4), 370–378. <https://doi.org/10.1177/1948550617693063>
- Funder, D. (2020). *Misgivings: Some thoughts about "measurement invariance."* <https://funderstorms.wordpress.com/2020/01/31/misgivings-some-thoughts-about-measurement-invariance/>
- Gebauer, J. E., Bleidorn, W., Gosling, S. D., Rentfrow, P. J., Lamb, M. E., & Potter, J. (2014). Cross-cultural variations in Big Five relationships with religiosity: A sociocultural motives perspective. *Journal of Personality and Social Psychology*, 107(6), 1064–1091. <https://doi.org/10.1037/a0037683>
- Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg University Press.
- Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes Jr., J. E. (1998). Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences*, 24(3), 393–403. [https://doi.org/10.1016/S0191-8869\(97\)00110-4](https://doi.org/10.1016/S0191-8869(97)00110-4)
- Gregorich, S. E. (2006). Do Self-Report Instruments Allow Meaningful Comparisons Across Diverse Population Groups? Testing Measurement Invariance Using the Confirmatory Factor Analysis Framework. *Medical Care*, 44(11), S78–S94. <https://doi.org/10.1097/01.mlr.0000245454.12228.8f>
- Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451–464. <http://doi.org/10.1016/j.labeco.2012.05.014>
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3), 411–482. <https://doi.org/10.1086/504455>
- Hussey, I., Hughes, S., & Nosek, B. A. (2018). The implicit and explicit [Data set]. In *Attitudes, Identities, and Individual Differences (AIID)*. <https://osf.io/pnjwf/>
- John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In *Handbook of Research Methods in Personality Psychology* (pp. 461–494).
- John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). Guilford Press.
- Jokela, M., & Keltikangas-Järvinen, L. (2011). The association between low socioeconomic status and depressive symptoms depends on temperament and personality traits. *Personality and Individual Differences*, 51(3), 302–308. <https://doi.org/10.1016/j.paid.2010.05.004>
- Jonassaint, C. R., Siegler, I. C., Barefoot, J. C., Edwards, C. L., & Williams, R. B. (2011). Low life course socioeconomic status (SES) is associated with negative NEO PI-R personality patterns. *International Journal of Behavioral Medicine*, 18(1), 13–21. <https://doi.org/10.1007/s12529-009-9069-x>
- Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52(3), 621–652. <https://doi.org/10.1111/j.1744-6570.1999.tb00174.x>
- Jung, E., & Yoon, M. (2017). Two-step approach to partial factorial invariance: Selecting a reference variable and identifying the source of noninvariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(1), 65–79. <https://doi.org/10.1080/10705511.2016.1251845>

- Kajonius, P. J., & Carlander, A. (2017). Who gets ahead in life? Personality traits and childhood background in economic success. *Journal of Economic Psychology, 59*, 164–170. <https://doi.org/10.1016/j.joe.2017.03.004>
- Kang, Y., McNeish, D. M., & Hancock, G. R. (2016). The Role of Measurement Quality on Practical Guidelines for Assessing Measurement and Structural Invariance. *Educational and Psychological Measurement, 76*(4), 533–561. <https://doi.org/10.1177/0013164415603764>
- Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychologist, 64*(4), 241–256. <https://doi.org/10.1037/a0015309>
- Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2016). Unemployment Alters the Set Point for Life Satisfaction. *Psychological Science, 15*(1), 8–13. <https://doi.org/10.1111/j.0963-7214.2004.01501002.x>
- Ludwig, R. M., Flournoy, J. C., & Berkman, E. T. (2019). Inequality in personality and temporal discounting across socioeconomic status? Assessing the evidence. *Journal of Research in Personality, 81*, 79–87. <https://doi.org/10.1016/j.jrp.2019.05.003>
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*(3), 471–491. <https://doi.org/10.1037/a0019227>
- McCrae, R. R., Costa Jr., P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., Sanz, J., Sánchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78*(1), 173–186. <https://doi.org/10.1037/0022-3514.78.1.173>
- McCrae, R. R., & Costa, P. T. (2008). Empirical and theoretical status of the five-factor model of personality traits. In *The SAGE Handbook of Personality Theory and Assessment* (Vol. 1, pp. 273–294). SAGE Publications Ltd. <https://doi.org/10.4135/9781849200462.n13>
- McCrae, R. R., Costa, P. T., de Lima, M. P., Simões, A., Ostendorf, F., Angleitner, A., Marušić, I., Bratko, D., Caprara, G. V., Barbaranelli, C., Chae, J.-H., & Piedmont, R. L. (1999). Age differences in personality across the adult life span: parallels in five cultures. *Developmental Psychology, 35*(2), 466–477. <https://doi.org/10.1037/0012-1649.35.2.466>
- Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116–130. <https://doi.org/10.1037/0022-3514.93.1.116>
- Nye, C. D., Bradburn, J., Olenick, J., Bialko, C., & Drasgow, F. (2019). How Big Are My Effects? Examining the Magnitude of Effect Sizes in Studies of Measurement Equivalence. *Organizational Research Methods, 22*(3), 678–709. <https://doi.org/10.1177/1094428118761122>
- Nye, C. D., & Drasgow, F. (2011). Effect Size Indices for Analyses of Measurement Equivalence: Understanding the Practical Importance of Differences Between Groups. *Journal of Applied Psychology, 96*(5), 966–980. <https://doi.org/10.1037/a0022955>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). <https://doi.org/10.1126/science.aac4716>
- Parker, K., & Stepler, R. (2017). *As US marriage rate hovers at 50%, education gap in marital status widens*. PEW Research Center.
- Penman-Aguilar, A., Carter, M., Snead, M. C., & Kourtis, A. P. (2013). Socioeconomic disadvantage as a social determinant of teen childbearing in the US. *Public Health Reports, 128*(2\_suppl1), 5–22. <https://doi.org/10.1177/00333549131282s102>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. <https://www.R-project.org/>
- Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality, 44*(1), 53–61. <https://doi.org/10.1016/j.jrp.2009.10.005>
- Revelle, W. (2018). *psych: Procedures for Personality and Psychological Research*. Northwestern University. <https://CRAN.R-project.org/package=psych>
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313–345. <https://doi.org/10.1111/j.1745-6916.2007.00047.x>
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>

- Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. *Handbook of Personality: Theory and Research*, 3, 375–398.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <http://www.jstatsoft.org/v48/i02/>
- Rosso, B. D., Dekas, K. H., & Wrzesniewski, A. (2010). On the meaning of work: A theoretical integration and review. *Research in Organizational Behavior*, 30, 91–127. <https://doi.org/10.1016/j.riob.2010.09.001>
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581–592. <https://doi.org/10.1093/biomet/63.3.581>
- Saegert, S., Adler, N., Bullock, H., Cauce, A. M., Liu, W. M., & Wyche, K. (2006). *FINAL REPORT APA TASK FORCE ON SOCIOECONOMIC STATUS (SES)*.
- Santiago, C. D., Wadsworth, M. E., & Stump, J. (2011). Socioeconomic status, neighborhood disadvantage, and poverty-related stress: Prospective effects on psychological syndromes among diverse low-income families. *Journal of Economic Psychology*, 32(2), 218–230. <https://doi.org/10.1016/j.joep.2009.10.008>
- Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, 64, 92–107. <https://doi.org/10.1037/h0022371>
- Schmitt, N., Golubovich, J., & Leong, F. T. (2011). Impact of measurement invariance on construct correlations, mean differences, and relations with external correlates: An illustrative example using Big Five and RIASEC measures. *Assessment*, 18(4), 412–427. <https://doi.org/10.1177/1073191110373223>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Shiner, R. L., Masten, A. S., & Roberts, J. M. (2003). Childhood personality foreshadows adult personality and life outcomes two decades later. *Journal of Personality*, 71(6), 1145–1170. <https://doi.org/10.1111/1467-6494.7106010>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113, 117–143. <https://doi.org/10.1037/pspp0000096>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, 100(2), 330–348. <https://doi.org/10.1037/a0021717>
- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, 101(4), 862–882. <https://doi.org/10.1037/a0024950>
- Srivastava, S. (2020). Personality structure: Who cares? *European Journal of Personality*, 54, 550–551. <https://doi.org/10.1002/per.2284>
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, 84(5), 1041–1053. <https://doi.org/10.1037/0022-3514.84.5.1041>
- Thalmayer, A. G., & Saucier, G. (2017). The questionnaire big six in 26 nations: Developing cross-culturally applicable big six, big five and big two inventories. *European Journal of Personality*, 28(5), 482–496. <https://doi.org/10.1002/per.2094>
- Vecchione, M., Alessandri, G., & Barbaranelli, C. (2012). The five factor model in personnel selection: Measurement equivalence between applicant and non-applicant groups. *Personality and Individual Differences*, 52(4), 503–508. <https://doi.org/10.1016/j.paid.2011.11.014>
- Wang, W. (2018, March 16). *Early Marriage Has Fallen, Especially Among Those Without a College Degree* [Blog post]. <https://ifstudies.org/blog/early-marriage-has-fallen-especially-among-those-without-a-college-degree>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Wood, D., Gardner, M. H., & Harms, P. D. (2015). How functionalist and process approaches to behavior can explain trait covariation. *Psychological Review*, 122(1), 84–111. <https://doi.org/10.1037/a0038423>
- Yoon, M., & Millsap, R. E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling*, 14(3), 435–463. <https://doi.org/10.1080/105510701301677>

Zisman, C., & Ganzach, Y. (2020). In a Representative Sample Grit Has a Negligible Effect on Educational and Economic Success Compared to Intelligence. *Social Psychological and Personality Science*, *12*(3), 296–303. <https://doi.org/10.1177/1948550620920531>

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*. <https://doi.org/10.1017/s0140525x17001972>

## SUPPLEMENTARY MATERIALS

### Peer Review History

Download: [https://collabra.scholasticahq.com/article/24431-the-big-five-across-socioeconomic-status-measurement-invariance-relationships-and-age-trends/attachment/61771.docx?auth\\_token=-dE3dz5WsXVStsjkDaFH](https://collabra.scholasticahq.com/article/24431-the-big-five-across-socioeconomic-status-measurement-invariance-relationships-and-age-trends/attachment/61771.docx?auth_token=-dE3dz5WsXVStsjkDaFH)

---

### Supplemental Material

Download: [https://collabra.scholasticahq.com/article/24431-the-big-five-across-socioeconomic-status-measurement-invariance-relationships-and-age-trends/attachment/61772.pdf?auth\\_token=-dE3dz5WsXVStsjkDaFH](https://collabra.scholasticahq.com/article/24431-the-big-five-across-socioeconomic-status-measurement-invariance-relationships-and-age-trends/attachment/61772.pdf?auth_token=-dE3dz5WsXVStsjkDaFH)

---