The Elusive General Factor of Personality:

The Acquaintance Effect

Timo Gnambs

Osnabrück University

Author Note

Timo Gnambs, Institute of Psychology, Osnabrück University, Germany.

Correspondence concerning this article should be addressed to Timo Gnambs, Institute of Psychology, Osnabrück University, Seminarstr. 20, 49069 Osnabrück, Germany, Email: timo.gnambs@uni-osnabrueck.de

Accepted for publication in *European Journal of Personality*

The definitive version of the article will be available at wileyonlinelibrary.com.

Abstract

A general factor (g_p) at the apex of personality has been suggested to account for the correlations between the Big Five. Although the g_p has received ample support from monomethod studies, results from studies incorporating different methods have remained rather ambiguous; some have identified a g_p across different informants whereas others have not. It was hypothesized that these divergent findings are a result of varying lengths of acquaintance between raters. To this end, the current study presents a multitrait multi-informant meta-analysis (total N = 11,941) that found weak support for a g_p as a substantive trait of personality. Evidence for a g_p was susceptible to the length of acquaintance between informants. Whereas a g_p could be identified for short-term acquaintances, it remained elusive at long-term acquaintance. Thus, the g_p in other ratings more likely reflects normative ratings of an average individual rather than ratings of the specific target person.

Keywords: Big Five, general factor, length of acquaintance, meta-analysis, multitrait multimethod

Word count: 149

The Elusive General Factor of Personality:

The Acquaintance Effect

Hierarchical views of personality (e.g., Carroll, 2002; Eysenck, 1947; Mowen, Park, & Zablah, 2007) describe personality as falling along a continuum that ranges from rather narrow traits to increasingly general dimensions. On the most abstract level, the Five-Factor Model (FFM; Digman, 1990) represents five orthogonal traits of personality: conscientiousness, agreeableness, neuroticism (or reverse scored as emotional stability), openness to experiences (or intellect), and extraversion. Although conceived as independent from each other, the five traits have routinely demonstrated low- to medium-sized correlations between their scores in empirical studies. Meta-analyses have estimated the mean true score correlation to be around $|\bar{r}| = .26$ to .29 (Mount, Barrick, Scullen, & Rounds, 2005; Rushton & Irwing, 2008). This has led some authors to speculate about a potential higher order hierarchy beyond the FFM (Digman, 1997; Musek, 2007). Despite receiving ample support in single-informant studies (e.g., Rushton & Irwing, 2008; Van der Linden, te Nijenhuis, Cremers, & Van de Ven, 2011), validity studies across multiple informants including self- and other ratings have been rather mixed. Some authors have identified a general factor of personality (Loehlin & Horn, 2012; Rushton et al., 2009), whereas others have not (Danay & Ziegler, 2011; Riemann & Kandler, 2010).

This paper seeks to explain these divergent findings as a result of varying levels of acquaintance between informants. The accuracy of observer ratings of personality frequently increases with the length of time they've known the target person (Beer & Watson, 2008a; Biesanz, West, & Millevoi, 2007; Kurtz & Sherker, 2003; Schneider, Schimmack, Petrican, & Walker, 2010). Therefore, if the general factor of personality represents a substantive trait of personality, it should be well-defined for long-acquainted individuals, whereas it is likely to emerge less clearly in dyads who have known each other for only a short period of time. On the other hand, if it primarily represents an evaluative bias resulting from stereotype information, the general factor would be expected to be better defined at short- than at long-term acquaintance. To this end, the current study presents a multitrait multimethod (MTMM) meta-analysis to study the effect of different levels of acquaintance on the emergence of a higher order general factor of personality across self- and other ratings.

Higher Order Models of Personality

A two-factorial view of personality postulates two orthogonal traits hierarchically superordinate to the five-factor space (Carroll, 2002; Digman, 1997): The α factor, also known as stability (DeYoung, Peterson, & Higgins, 2002, 2005), represents low levels of neuroticism and high levels of conscientiousness and agreeableness, whereas the β factor (or plasticity) reflects the shared variance between openness and extraversion. These two *superfactors* (Carroll, 2002) or *metatraits* (Digman, 1997) have been suggested to reflect individual differences in self-control and personal growth as seen in the restraint of hostile and aggressive behaviors toward others and an active engagement with the environment (Hirsh, DeYoung, & Peterson, 2009). They express two fundamental needs of individuals: the need for stable psychosocial functioning and the need for an active exploration of the world (DeYoung et al., 2002, 2005). Together, they determine how individuals react in novel situations. These metatraits loosely resemble Block's (Block & Block, 1980; see also Robins, John, & Caspi, 1994) two-factorial personality model that has been introduced as an early alternative to the FFM and describes two central traits, ego-control and ego-resilience. The former refers to the capacity to inhibit one's impulses and, thus, mimics stability, whereas the latter determines the capacity to adapt one's reaction to situational demands. Support for the two-factorial structure of personality has been received from several single sample studies (e.g., Alessandri & Vecchione, 2012; Hirsh et al., 2009) and also various multimethod examinations (e.g., DeYoung, 2006; McCrae et al., 2008; Şimşek, Koydemir, & Schütz, 2012). Although there is still some debate if both factors are equally pronounced across cultures—for example, some European and Asian studies could not univocally confirm the α factor (cf. Jang et al., 2006)—overall, meta-analytical summaries clearly reproduced both factors (Chang, Connelly, & Geeza, 2012; Markon, Krueger, & Watson, 2005). These factor-analytical studies combined with accumulated evidence of a neurobiological basis of the two metatraits (DeYoung et al., 2002; DeYoung, Hasher, Djikic, Criger, & Peterson, 2007) led Block (2010) to conclude in his review that the five factors of personality are clearly "subsumed by the higher order, progenetive Big Two factors" (p. 21).

The general factor of personality, g_p (Musek, 2007), represents the most abstract level of personality and is assumed to be hierarchically superordinate not only to the FFM but also to the two-factor model of personality. It constitutes a combination of those Big Five components that are generally positively valued: high levels of openness, conscientiousness, extraversion, and agreeableness and low levels of neuroticism. High scorers on the g_p have been attributed a "good" personality (Rushton & Irwing, 2011, p. 132) and are seen as friendly, well-adjusted, and outgoing, whereas low scorers are characterized as "difficult" personalities that don't mix well with others. In this respect, the g_p has been associated with various favorable characteristics such as positive affectivity, subjective well-being (Musek, 2007), self-esteem (Erdle, Irwing, Rushton, & Park, 2010), and even general intelligence (Loehlin, 2011). Moreover, the validity of the g_p has been inferred from its prediction of various behavioral outcomes. For example, the g_p predicted job performance of long-term employees in business organizations and military personnel (Van der Linden et al., 2011). In adolescents, it was related to sociometric position within the peer group and ratings of likability (Van der Linden, Scholte, Cillessen, te Nijenhuis & Seggers, 2010).

The g_p has been recovered in various single-method studies in mixed samples of the general population (Erdle et al., 2010), children (Van der Linden et al., 2010), and even psychiatric patients (Van der Linden, te Nijenhuis, & Bakker, 2010). However, monomethod studies are distorted to some degree because true trait components cannot be distinguished from rater-specific biases, for example, a self-favoring bias (Paulhus, Bruce, & Trapnell,

1995) that leads to inflated ratings of one's standing on a particular trait. In particular, self-reports are prone to a common method bias (Podsakoff, McKenzie, Lee, & Podsakoff, 2003), which results in spurious correlations between measures of different constructs obtained from the same source. This seems particularly relevant for the case of a general factor of personality. Although some research has identified a g_p across self- and peer reports as well (Loehlin & Horn, 2012; Rushton et al., 2009), others have not (Danay & Ziegler, 2011; Riemann & Kandler, 2010); a recent meta-analysis found only weak support for a g_p across multiple informants (Chang et al., 2012). For example, Anuisc, Schimmack, Pinkus, and Lockwood (2009) suggested that the g_p is a product of informant-specific Halo error reflecting a general disposition to attribute favorable characteristics to oneself and others. An explanation for the mixed support of the general factor hierarchy in multi-informant studies might be attributed to varying levels of acquaintance within the rater dyads.

The Effect of Acquaintance Length

Acquaintance between two individuals refers to the degree to which they are familiar with or have knowledge about each other. It is comprised of qualitative (i.e., type of relationship) and quantitative (i.e., frequency and intensity of interactions) aspects (Starzyk, Holden, Fabrigar, & MacDonald, 2006). The accuracy of trait judgments is frequently a function of the quantitative aspect: the length of acquaintance (Biesanz et al., 2007; Bernieri, Zuckerman, Koestner, & Rosenthal, 1994; Kurtz & Sherker, 2003; Paulhus & Bruce, 1992; Schneider et al., 2010). Long-term acquaintances have more opportunities to interact with each other and observe each other's behaviors in different situations, and this typically makes them better informants than short-term acquaintances. For example, Watson, Hubbard, and Wiese (2000) observed that the agreement between self- and other-reported personality is about $\Delta r = .15$ higher for long-wed couples than respective correlations for short-term friend dyads. Further support for this acquaintanceship effect has been received from longitudinal studies that have demonstrated increasing self-other agreement over time. Paulhus and Bruce

(1992) examined agreement within initially unacquainted groups that met each other over the course of 7 weeks. Agreement between self- and informant ratings of personality increased significantly over time. A similar trend was identified in pairs of college roommates over a period of 4 month (Kurtz & Sherker, 2003). Biesanz et al. (2007) estimated an increase in self-other agreement of about $\Delta r = .05$ for every 5 years of acquaintance, whereas other authors (Schneider et al., 2010) believe that the accuracy of trait ratings monotonically increases during only the first 3 years of acquaintance; beyond that, length of acquaintance does not ensure higher self-other agreement. This effect is typically more pronounced for those traits in the five-factor space that are less clearly manifested in observable behaviors (i.e., neuroticism, agreeableness, or openness; Kurtz & Sherker, 2003; Paulhus & Bruce, 1992; Simms, Zelazny, Yam, & Gros, 2010). By contrast, extraversion and conscientiousness, which are even readily inferred from thin slices of behavior (e.g., Carney, Colvin, & Hall, 2007) show high levels of self-other agreement early on in a relationship, and this agreement shows little increase over time (Paulhus & Bruce, 1992; Simms et al., 2010).

The effect of acquaintance length has been attributed to differential effects of stereotype (Cronbach, 1955) or normative (Furr, 2008) information about what people generally tend to be like. If substantial information about an individual’s trait level is not available, peers resort to implicit personality theories, a set of preexisting beliefs about people and how traits typically covary, and substitute missing information with stereotypical estimates of the “average” or “typical” person’s trait (Beer & Watson, 2008a). These a priori beliefs function as a form of heuristic to simplify personality ratings made by others and to create a coherent personality impression. The stronger this simplicity heuristic, the less accurately people distinguish between different personality dimensions and, thus, cluster different traits along a common continuum. Because normative ratings are generally rather positive in nature (Wood, Gosling, & Potter, 2007), making observer ratings at short-term acquaintance also entails viewing others very positively. Consequently, these ratings by others result in an attenuation or even denial of socially undesirable attributes, and this could lead to trait judgments that resemble a general factor of personality.

If the general factor beyond the five-factor space is not merely an artifactual bias in self- or other perceptions but a substantive structure of personality, it should be unaffected by the length of acquaintance. On the other hand, if the higher order structure fails to replicate at long-term acquaintance and can only be identified at short-term acquaintance, it is more likely to be a product of stereotype-based judgments. These stereotype effects should result in higher cross-informant correlations for similar positively evaluated traits and, thus, artificially create a general factor of personality.

Overview

The higher order structure of personality was analyzed in a meta-analysis of multi-informant correlations of the five factors of personality assessed as self- and peer reports. The study reconstructed a full multitrait multi-informant matrix consisting of correlations between the Big Five resulting from self- and other ratings. For each correlation in this matrix, a separate meta-analysis was conducted, thus resulting in 45 independent meta-analyses. In the second step, the synthesized correlations were analyzed in search of a general factor of personality. Then the length of acquaintance between the raters was considered as a potential factor that might mask the identification of a higher order structure in the synthesized multi-informant data.

Method

Literature Search

Primary studies reporting relevant correlations between measures of the Big Five obtained from self and nonself sources were located by searching several computerized databases (PsycINFO, Psyndex, EconLit, and Google Scholar) using the keywords “(trait or Big Five or Five Factor Model) and (peer or informant or observer or spouse or roommate or self-other).” Moreover, references of previous meta-analyses on self-other agreement of personality (Chang et al., 2012; Connolly, Kavanagh, & Viswesvaran, 2007; Connelly & Ones, 2010; McCrae et al., 2004) and the manuals of published personality inventories were inspected for additional studies reporting self-other correlations of personality.

A study was included in the meta-analysis when it met the following criteria: (a) The study was published after 1980,[1] (b) it was written in English or German, and (c) it included a measure of personality according to the five factor taxonomy. Eligible Big Five instruments were identified using the classification by Salgado (2003). Instruments not included in this classification were categorized as Big Five measures based on the evaluations of two independent raters. To avoid artifactual errors due to imperfect construct validities (cf. Hunter & Schmidt, 2004; Mount & Barrick, 1995), instruments that were developed outside the five-factor framework were excluded. (d) The traits were measured with a validated multi-item instrument. Scales that were constructed ad hoc or single-item measures were excluded to avoid spurious correlations resulting from unreliable instruments. (e) Personality ratings of at least one of the five traits were obtained from other ratings. (f) The study reported correlations between traits measured by the same informant or cross-informant agreement. Studies reporting profile analyses or mean differences[2] were excluded. (g) The mean duration of the acquaintance between the target person and the observer was reported. (h) Participants, raters, and ratees were at least 14 years of age and (i) of sound physical and psychological health. Studies on children or patients with severe physical trauma or mental illnesses were not considered in order to exclude individuals with unstable personalities for whom temporary personality changes seemed likely.

---

[1] This marks the time Goldberg (1981) coined the term “Big Five” and wide-spread acceptance of the Five-Factor Model as a broad taxonomy of human personality began to emerge (cf. John, Naumann, & Soto, 2008).

[2] Although standardized mean difference scores can be transformed into correlation coefficients, empirical evidence suggests that the two effect size measures are largely independent from each other (Fletcher & Kerr, 2010).

This search resulted in 44 eligible research articles and three theses reporting 1,481 correlation coefficients.

Meta-Analytic Procedure

In order to identify higher order factors of personality from the multi-informant data, in the first step, a 10 x 10 matrix was formulated consisting of true-score correlations between (a) the five self-reported personality traits, (b) the five peer-reported traits, and (c) the five traits assessed by different raters. For each correlation in this matrix, a separate meta-analysis was conducted, thus resulting in 45 independent meta-analyses.

Nonindependence. Untransformed Pearson product moment correlations were used as effect size measures. To ensure an appropriate level of independence, the following approaches were used: (a) For studies reporting on several independent samples, correlations from each sample were included; (b) When studies reported multiple correlations for the total sample and several subgroups, only the total sample correlation was considered; (c) If a study included multiple correlations between two traits from the same sample (e.g., measured with different instruments), the correlations were combined into a composite correlation using the procedure proposed by Cheung and Chan (2004). This resulted in 986 independent correlation coefficients from 56 samples.

Outliers. Extreme correlations (i.e., outliers) were identified using the studentized deleted residual (Viechtbauer & Cheung, 2010), which yields a *z*-standardized difference measure between each observed effect and the predicted average true effect when the respective effect actually fits the assumed model. Using a nominal α of 1%, these indicated that between 0 and 2 correlations were potential outliers. To reduce the impact of these outliers, the respective correlations were truncated to the lower or upper bound of the 90%

credibility interval of the true effect calculated from a dataset from which the outliers had been removed.[3]

Effect size synthesis. Correlations were synthesized using a random effects model with a restricted maximum likelihood estimator (Viechtbauer, 2005), which decomposes the variability of the effect sizes into heterogeneity due to random population effects and sampling variance. In contrast to fixed-effect models, these models do not assume an identical population parameter across all studies—which is seldom tenable in empirical research synthesis (see Schmidt, Oh, and Hayes, 2009, for a review). The accuracy and significance of the synthesized effects were gauged by means of a 95% credibility interval.

Correction for artifacts. The observed correlations were corrected for two sources of error: sampling error and measurement error. Sampling error was accounted for by weighing the individual correlations by the inverse of their variances. Measurement error was accounted for twofold. First, since some studies employed multiple peer informants which are likely to result in higher reliabilities than ratings from a single informant these correlations were individually corrected using the interrater reliabilities following the approach in Chang et al. (2012)[4]. Second, adjustments for the instruments' test-retest reliabilities were applied. These corrected correlations represent the stable overlap between self- and other ratings with situation-specific random variance from differences in, for example, mood or alertness removed (Connelly & Ones, 2010). As none of the primary studies reported information on test-retest reliabilities, a separate meta-analysis on test-retest correlations for personality

---

[3] Sensitivity analyses that did not account for these extreme correlations resulted in slightly larger random variance components of the synthesized correlations, but did not yield different results regarding the multitrait multi-informant analyses.

[4] Each correlation was individually disattenuated for the interrater reliability for multiple raters reported in the study and subsequently reattenuated for the reliability of a single rater using the meta-analytically derived reliability from Connelly and Ones (2010).

inventories assessing the Big Five was conducted.[5] The means and variances of the square roots of these synthesized test-retest correlations were used as artifact distributions to correct the variance-weighted mean correlations for transient error (Hunter & Schmidt, 2004). Other forms of measurement error such as internal consistency were not considered as these hardly affect self-other correlations (McCrae, Kurtz, Yamagata, & Terraciano, 2011).

Multitrait-Multimethod (MTMM) Analyses

Latent variable modeling. The correlations between the Big Five synthesized in the first step were subjected to structural equation modeling (SEM; cf. Cheung & Chan, 2005; Viswesvaran & Ones, 1995) in Mplus 6 (Muthén & Muthén, 1998-2011) with a maximum likelihood estimator. Following recommendations by Viswesvaran and Ones (1995), the harmonic mean of all samples was used as the sample size for these analyses because the harmonic mean gives less weight to individual large studies than the arithmetic mean and, as such, more closely reflects the overall precision of the data. The choice of sample size in meta-analytic SEM primarily affects the parameters' standard errors (and consequently the associated significance tests), but not the parameter estimates themselves.

MTMM models. All analyses modeled five latent trait factors, each represented by two indicators: the self-rating and the peer rating. Thus, each latent trait represented the variance shared across informants. To identify the latent factors, the paths for the two indicators were constrained to be equal; thus, self- and peer ratings contributed equally to the latent trait variance. First, a baseline model was specified that included five correlated traits without acknowledging informant-specific biases. This model was subsequently extended

---

[5] The artifact distributions had the following means and standard deviations of the square root of test-retest correlations for Big Five instruments administered twice within a period of at most 8 weeks: conscientiousness (M = .92, SD = .03, k = 130, N = 13,011), agreeableness (M = .89, SD = .04, k = 100, N = 13,705), neuroticism (M = .91, SD = .03, k = 158, N = 14,103), openness (M = .91, SD = .02, k = 116, N = 12,274), and extraversion (M = .93, SD = .02, k = 145, N = 14,351).

with correlated error terms for each informant to acknowledge rater-specific biases. Then a higher order trait model that included two correlated higher order trait factors, α and β, was tested (see left panel of Figure 1). To identify the β factor, the loadings of its indicators were constrained to be equal. Finally, to separate the α and β factors from a potential general factor of personality, a bifactorial model with a general factor in addition to two orthogonal α and β factors was considered (see right panel of Figure 1). As bifactorial models with five traits are ordinarily not identified, the respective factor loadings were estimated using the Schmid-Leiman (1957) procedure. This transforms the oblique factor structure obtained in the correlated α and β model into a bifactor structure with a common factor (in this case, the general factor of personality) and two orthogonal group-specific factors (cf. Reise, 2012). As a global indicator of the general factor's importance, McDonald's (1999) ω_h was reported. ω_h represents the ratio of variance accounted for by the general factor to the total amount of variance explained by all factors and has been suggested to be an optimal indicator of a measure's general factor saturation (Zinbarg, Revelle, Yovel, & Li, 2005). As a simple rule-of-thumb (Revelle, 1979), a ω_h of at least .50 has been suggested as a minimum threshold in order to allow for meaningful interpretations of the common factor.

Model evaluation. Model fit was evaluated in line with common praxis (cf. Schermelleh-Engel, Moosbrugger, & Müller, 2003) using the comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Different models were compared with the sample-size-adjusted Bayesian Information Criterion (BIC) for which lower values indicate a better fit.

Results

Study Characteristics

Most samples originated from North America (63%) and Europe (33%). The total sample size was N = 11,941 (range of the individual studies' Ns: 33 to 1,260), and approximately 61% of the participants were female. Ages ranged from 14 to 63 years (M =

29.43, *SD* = 12.91). The type of relationship between target and informant was qualified as relative (e.g., parent, sibling) for 12%, spouse or dating partner for 24%, friend or close acquaintance for 16%, incidental acquaintance or stranger for 16%, and unspecified peer for the remaining dyads. The length of acquaintance between the raters ranged from less than a year to 35.5 years (*Mdn* = 4 years). For most samples, other ratings were based on a single informant; about 15% included two informants and 10% up to nine. Most studies used variants of Costa and McCrae's (1992) NEO scales (43%), followed by various adjectives lists (31%), the Big Five Inventory (12%; John et al., 2008), and Goldberg's (1999) statements from the International Personality Item Pool (2%).

Synthesized Correlations

In total, 45 separate meta-analyses were conducted, one for each correlation resulting from the assessment of the five factors of personality as self- and other reports. The results of these meta-analyses are summarized in Table 1. All meta-analyses involved between 16 and 52 independent effect sizes based on a minimum of *N* = 4,337 participants. Most effect sizes were available for the syntheses of self-other correlations of the same trait (range: 47 to 52). The corrected self-other correlations for all five traits (see bold values in Table 1) demonstrated good convergent validities across raters, with all values falling between .43 (neuroticism) and .59 (extraversion). These results are comparable to self-other correlations obtained in previous meta-analyses (cf. Connelly & Ones, 2010; Connolly et al., 2007). Most heterotrait-heteromethod correlations were small ($r < .15$) and not significant ($p > .05$). Within informants, the five traits were moderately correlated: $|\bar{r}| = .21$ for self-ratings and $|\bar{r}| = .26$ for peer ratings.

Meta-Analytic MTMM Analyses

The synthesized 10 x 10 matrix of true-score correlations was used as input for the confirmatory factor analyses in the search for higher order factors of personality. A model with five correlated latent trait factors, but without informant-specific biases did not provide an adequate fit to the data, $\chi^2(30) = 6{,}254$, CFI = .597, RMSEA = .204, SRMR = .092, BIC = 133,068. Acknowledging potential method effects by modeling correlated error terms in addition to the five traits achieved a superior fit, $\chi^2(10) = 52$, CFI = .997, RMSEA = .029, SRMR = .012, BIC = 127,037. The mean absolute correlations between the latent traits was .12, which is similar to correlations obtained in previous single-sample multitrait-multimethod studies (e.g., DeYoung, 2006: $|\bar{r}| = .11$ and .15; Riemann & Kandler, 2010: $|\bar{r}| = .07$). Thus, even across multiple perspectives, when informant-specific variations were removed, the five traits were not completely orthogonal and rather remained significantly correlated. The pattern of these correlations appeared to be consistent with the implied structure of two higher order factors (Digman, 1997) resulting from a significant ($p < .001$) correlation between openness and extraversion ($r_{OE} = .14$) on the one hand, and from correlations between neuroticism, conscientiousness, and agreeableness on the other hand ($r_{NC} = -.21$, $r_{NA} = -.19$, and $r_{CA} = .12$, respectively).

The higher order factor model with two correlated α and β factors, $\chi^2(15) = 238$, CFI = .986, RMSEA = .054, SRMR = .041, BIC = 127,180, resulted in a satisfactory fit. Although all traits had significant loadings on their higher order factor, the α factor was primarily defined by neuroticism ($\lambda = .70$, $p < .001$, $R^2 = .50$), whereas agreeableness ($\lambda = .30$, $p < .001$, $R^2 = .09$) and conscientiousness ($\lambda = .31$, $p < .001$, $R^2 = .10$) had somewhat moderate loadings (see left panel of Figure 1). The two metatraits were significantly correlated at $r = .33$, $p < .001$, which falls in line with the assumption of a superordinate general factor of personality at another level above the α and β factors.[6] The estimates in the bifactor model resulted in

---

[6] An explicit modelling of the g_p does not provide additional information as compared to the bivariate correlation because a higher order factor with only two indicators is undetermined. A typical solution is to constrain the loadings of the two indicators to be equal. As a result, the standardized factor loadings on the g_p correspond to the square root of the correlation (here: $\lambda = \sqrt{.33} = .57$).

moderate loadings on the common factor for most traits (see right panel of Figure 1) between $\lambda = .17$ for agreeableness and $\lambda = .40$ for neuroticism. As a consequence, McDonald's (1999) ω_h, an indicator of the general factor saturation, was rather low ($\omega_h = .21$). As a common factor should at least explain 50% of the variance in order to allow for meaningful interpretations (Revelle, 1979), evidence of a general factor of personality common to all five traits was rather scarce in the cross-informant correlations.

Length of Acquaintance

A potential higher order structure of personality could have been masked by varying levels of acquaintance. The acquaintance effect on the emergence of a g_p was studied in two ways: On the one hand, short-term acquaintances were compared to long-term acquaintances by means of subgroup analyses. On the other hand, gradients of the focal parameters across different levels of acquaintance were analyzed using local weights for each individual effect size (see online supplement for more details about the procedure).

Subgroup analyses. The available samples were split into two subgroups. Because the validity of peer ratings increases markedly within the first 3 years of acquaintance, but increases less beyond that point (Schneider et al., 2010), the short-term acquaintance group included samples with a median acquaintance length of *Mdn* = 6 months (range = [0, 36]). By contrast, long-term acquaintances knew each other on average *Mdn* = 13.77 years (range = [3.42, 35.5]). In line with previous observations (e.g., Biesanz et al., 2007; Kurtz & Sherker, 2003; Paulhus & Bruce, 1992), self-other agreement increased at long-term acquaintance (see Table 2). This was most pronounced for openness, $\Delta\bar{r} = .22$, $p < .001$, neuroticism, $\Delta\bar{r} = .17$, $p < .001$, agreeableness, $\Delta\bar{r} = .15$, $p < .001$, and conscientiousness, $\Delta\bar{r} = .13$, $p < .001$; but to some degree also for extraversion, $\Delta\bar{r} = .08$, $p < .001$. However, the most striking difference resulted for the latent heterotrait correlations, which, in most cases, were smaller at long-term acquaintance; for example, the difference in correlations between agreeableness and conscientiousness was $\Delta r = -.17$, $p < .001$, and the difference between conscientiousness and neuroticism was $\Delta r = -.11$, $p < .001$. The only exception was the correlation between openness and extraversion which increased by $\Delta r = .12$, $p < .001$ (see Table 3).

As a consequence, higher order latent trait models resulted in a better fit at short-term acquaintance, $\chi^2(15) = 97$, CFI = .977, RMSEA = .059, SRMR = .042, BIC = 40,917, than at long-term acquaintance[7], $\chi^2(16) = 155$, CFI = .990, RMSEA = .051, SRMR = .042, BIC = 81,425. At short-term acquaintance, all traits had significant loadings on their respective higher order factors. Moreover, the α and β factors were highly correlated ($r = .57$, $p < .001$), and the bifactor estimates (see bottom left panel of Figure 2) also included substantial loadings on a common factor. By contrast, at long-term acquaintance, the respective higher order factor structure was markedly different (see top right panel of Figure 2). The correlation between α and β dropped to $r = .24$, $p < .001$. As a result, the common factor was rather ill defined (see bottom right panel of Figure 2). The g_p predominantly represented neuroticism ($R^2 = .28$) and explained only between 0 – 3% of the variance of the other traits. The different loading pattern on the g_p was mirrored by ω_h, which was higher at short-term acquaintance ($\omega_h = .38$) than at long-term acquaintance ($\omega_h = .18$).

Gradients of model parameters. The subgroup analyses presented in the previous section had two limitations: First, the varying lengths of acquaintance within each subgroup were ignored. Second, the chosen length of acquaintance used to divide the samples into the short- and long-term acquaintance groups was arbitrary to some degree. To overcome these limitations, a cross-sectional gradient for the latent correlation between the two higher order factors α and β was estimated (cf. Hildebrandt, Sommer, Herzmann, & Wilhelm, 2010). At

---

[7] At long-term acquaintance, the higher order model initially resulted in a minor negative residual variance for neuroticism, $\sigma^2 = -.10$ (see DeYoung, 2006, for a comparable problem). Although this might indicate structural misspecification, more often it is a result of a true parameter of or close to zero (Chen, Bollen, Paxton, Curran, & Kirby, 2001). A Wald test confirmed that the residual variance for neuroticism did not differ significantly from zero, $\chi^2(1) = 0.067$, $p = .80$. Thus, at long-term acquaintance the respective residual was fixed to zero.

focal points from 0 to 20 years of acquaintance, the previous meta-analyses and subsequent latent variable models were estimated anew using local weights for the individual effect sizes. The weights were created in such a way that effect sizes from samples near the defined focal point were given a larger weight approaching 1, whereas effect sizes from samples distant from the focal point were given smaller weights approaching 0 (see the online supplement for more details). Hence, the previous analyses were repeated 21 times using different weights depending on the focal length of acquaintance. This allowed for the inspection of continuous parameter changes across different lengths of acquaintance without creating a priori subgroups. The loadings of the five traits on the two higher order factors, α and β, across different lengths of acquaintance are plotted in the left panel of Figure 3. Most traits showed a gradual increase of their factor loadings with long-term acquaintance; only agreeableness demonstrated a marked drop in factor loadings. The latent correlations between the two higher order factors across different lengths of acquaintance are plotted in the left panel of Figure 4. In line with the previous subgroup analyses, the g_p, as indicated by the correlation between α and β, was most pronounced among short-term acquaintances (e.g., at 1 year, it was $r = .54$, $p < .001$). After 20 years of acquaintance, the respective correlation dropped to $r = .11$, $p = .34$. This decline was mirrored by McDonald's ω_h; the g_p emerged more clearly at 1 year of acquaintance ($\omega_h = .32$), whereas it was increasingly difficult to identify for individuals who had known each other for a longer period of time ($\omega_h = .08$).

Sensitivity analyses. Previous studies indicated mixed support of the higher order structure of personality across different cultural regions (e.g., Jang et al., 2006). Therefore, the previous analyses were repeated for a subgroup of samples that were conducted in North America (US and Canada). Among these samples the previously reported results were clearly confirmed. Self-other correlations significantly, $p < .05$, increased with length of acquaintance (see Table S1 of the online supplement) whereas most cross-trait correlations decreased (see Table S2). Moreover, acquaintanceship length moderated the emergence of the g_p. The correlation between α and β was strongest among short-term acquaintances and gradually decreased within 20 years of acquaintance (see right panel of Figure 4). Among European samples higher order models failed to converge because agreeableness was uncorrelated to conscientiousness, $r = .05$, $p = .19$, and neuroticism, $r = .03$, $p = .44$. As a consequence, neither the α factor nor a putative g_p could be identified. However, these results should be interpreted with caution because they are based on rather few primary studies—many of the meta-analyses conducted to construct the correlation matrix for the confirmatory models included as few as six primary studies.

Discussion

The observation that the Big Five are empirically frequently moderately correlated (e.g., Mount et al., 2005) has led to the proposal of a general factor at the apex of the personality hierarchy, similar to the cognitive domain (Musek, 2007; Rushton & Irwing, 2011). However, previous validity studies across multiple informants have yielded rather mixed results; some studies identified the g_p (e.g., Rushton et al., 2009), whereas others did not (e.g., Riemann & Kandler, 2010). In order to rectify these seemingly contradictory findings, the present study reported a multitrait-multimethod analysis on meta-analyzed self- and peer reports of personality. By extending recent meta-analytical research on the structure of personality (Chang et al., 2012) with findings on implicit simplicity effects in observer ratings of personality (Beer & Watson, 2008a, 2008b) the study provided several new insights: (a) Even when controlling for method effects, the Big Five are moderately correlated. Although a putative g_p could be extracted from these correlations, the respective factor loadings on the five traits were rather small. (b) The length of acquaintance between the informants moderated the identification of a g_p. In line with an artifact interpretation of the g_p, the common factor emerged more clearly at short- than at long-term acquaintance. (c) The two-factorial higher order structure was less susceptible to the length of acquaintance between informants. If anything, the factor loadings of most traits (except for agreeableness) on α and β tended to increase with acquaintanceship length. (d) The view of a universal higher order structure of personality across cultures might be challenged: the α factor could not be identified among European countries whereas it clearly emerged in North American samples. (e) Finally, the presented study also offered a methodological contribution by presenting a new method for the analysis of continuous moderators in meta-analytical SEM using parameter gradients (see online supplement).

The g_p Across Informants

Personality research is dominated by single-method studies, which cannot adequately separate true trait components from artifacts that are a result of a specific measurement method. Unfortunately, the bulk of previous research supporting the existence of a g_p has relied on monomethod studies (e.g., Erdle et al., 2010; Rushton & Irwing, 2008; Van der Linden et al., 2011). As soon as the higher order structure of personality was examined across multiple instruments (e.g., Hopwood, Wright, & Donnellan, 2011) or informants (e.g., Danay & Ziegler, 2011), the previously impressive apparent support for a g_p became less clear. In line with previous single-sample studies that managed to successfully identify a g_p across different raters (e.g., Loehlin & Horn, 2012; Rushton et al., 2009), the present multitrait multi-informant meta-analysis identified moderate correlations within the five factor space that allowed for the extraction of a putative g_p. Although the respective factor loadings on the g_p (λ = .57) fell in line with previous monomethod studies (λ = .63 - .67; Rushton & Irwing, 2008), the g_p explained only about 3 - 16% of the variance in the Big Five. Thus, one might question the meaningfulness of such a trait beyond the Big Five. The interpretation of the g_p as a substantive personality trait is further challenged by its susceptibility to the length of acquaintance between informants. Although self-other agreement in the current study increased with length of acquaintance—thus, replicating previous findings (cf. Biesanz et al., 2007; Watson et al., 2000)—the respective cross-trait correlations were significantly reduced in size (about $\Delta|\bar{r}| = .10$). As a consequence, a g_p was identified at short-term acquaintance, whereas it emerged less clearly for long-term acquaintances. This lends less plausibility to viewing the g_p as a substantive trait of personality and rather sustains an artifact interpretation.

The g_p as Shared Bias

If the g_p represents a bias in self- and other reports, how might this bias be explained? Observers who have not known a target person long enough to have sufficient information regarding his or her actual personality typically substitute missing information with normative information on how people typically are or ought to be (Beer & Watson, 2008a). In such a way, they try to create a consistent personality image of others. In support of this premise, Biesanz et al. (2007) demonstrated that observers' ratings of a target person reflect the ratings of an average hypothetical target instead of the specific individual when the length of acquaintance is short. Moreover, zero-acquaintance studies using photos have revealed moderate correlations between socially desirable characteristics; that is, faces rated high in agreeableness, which represents the most socially favorable attribute of the Big Five (Hafdahl, Panter, Gramzow, Sedikides, & Insko, 2000), are also rated as being somewhat high in openness, conscientiousness, extraversion, and emotional stability (Penton-Voak, Pound, Little, & Perrett, 2006). However, with increasing acquaintance, this spill-over effect becomes smaller. The longer observers have known the target person, the more traits become differentiated and stereotypical ratings shrink in favor of ratings that more closely reflect the respective individual. For example, the mean intercorrelation between other ratings of the Big Five decreased about $\Delta r = .22$ for ratings of complete strangers as compared to ratings of spouses (Beer & Watson, 2008b). Thus, at short-term acquaintance, others are attributed a variety of socially favorable qualities such as being agreeable, intellectually curious, and emotionally stable, which together mimic a putative g_p. A rather similar effect can be observed in self-ratings of personality. Meta-analytical reviews (e.g., Li & Bagger, 2006) have associated all traits within the FFM with socially desirable responding, the tendency to present oneself overly favorably in line with prevalent social norms (cf. also the *Halo effect*;

Anusic et al., 2009). Typically, it leads to inflated ratings on all five traits within the FFM (Paulhus et al., 1995). Social desirability is frequently captured by the first principal component of self-report inventories (Edwards & Edwards, 1991; Schmitt & Ryan, 1993) and, as a consequence, accounts for a significant proportion of variance in the g_p (e.g., Backström, 2007; Musek, 2007). This effect is partly a consequence of the evaluative item content of most FFM instruments. When neutrally rephrased items are administered from which the socially desirable content has been removed, evidence for a g_p gradually disappears (Backström et al., 2009). Thus, social desirability results in a bias in self-ratings similar to that found for other ratings at short-term acquaintance. Because the bias in other ratings gradually decreases the longer observers have known the target person, in the present meta-analysis, the g_p across informants became less evident with increasing length of acquaintance. After about 10 years of acquaintance, the g_p gradually disappeared (see Figure 4). In line with an artifact interpretation, these analyses indicate that a g_p that converges across raters (cf. Rushton et al., 2009) is primarily the result of normative information in other reports of personality.

The g_p as More Than Bias?

Following the tradition of Campbell and Fiske (1956) this study examined the g_p from a multiple informant perspective. Because single-method studies cannot separate true trait components from artifacts resulting from the measurement method, multi-method studies have been frequently advocated for the validation of constructs in the personality domain (cf. Schimmack, 2010). These analyses examined the variance shared across self- and other ratings to identify a putative g_p; unshared variance components unique to the self or the other perspective were treated as measurement error. However, it is conceivable that these unshared variance components not only represented error but also included substantial aspects of an individual’s personality (cf. Vazire & Carlson, 2011). Self-reports might contain information about oneself that is not readily observable by others, just as observer reports might include information about another person’s personality that goes unnoticed by oneself. In line with this assumption Vazire and Mehl (2008) demonstrated that self- and observer reports of typical behaviors differentially predicted a person's actual behaviors; the self was more accurate at predicting some behaviors whereas observers were more accurate at others. Thus, each informant had access to specific information not available to the other. This is also highlighted by several criterion validity studies of the FFM traits. A recent meta-analysis demonstrated that observer ratings of personality predicted job performance and showed incremental validity beyond self-reports (Oh, Wang, & Mount, 2011). Thus, the observer ratings included specific information about a person's personality not captured by the respective self-ratings and uniquely predicted work behaviors. Similarly, implicit aspects of personality that are not readily accessible to oneself but can manifest in spontaneous behaviors observable by others predicted actual behaviors beyond explicit trait ratings (Back, Schmukle, & Egloff, 2009). In light of several single-informant studies demonstrating the criterion validity of the g_p (e.g., Van der Linden et al., 2010, 2011) it might be speculated that some of the informant-specific variance in this study included substantial trait components that are not shared across perspectives. This would leave some room for a rater-specific g_p beyond a mere bias interpretation that should be explored in future studies.

A Universal Higher Order Structure of Personality

The g_p is but one recent attempt to pattern the correlations between the five traits of personality. Although the current study provided scarce evidence for a g_p that replicates across different lengths of acquaintance and cultures, the two-dimensional structure proofed to be more robust. Plasticity, the correlation between extraversion and openness, emerged clearly in short- and also long-term acquaintance groups—although the factor loadings on the higher order trait tended to increase gradually with increasing length of acquaintance. Thus, plasticity seemed to be better defined in pairs that knew each other a longer time. Moreover, plasticity was also the only higher order factor that replicated across cultures; the β factor emerged comparably in North American and European samples. Stability, the second higher order factor of personality, was also clearly identifiable albeit not invariant across different lengths of acquaintance (see Figure 3): the loadings for neuroticism and conscientiousness gradually increased for long-time acquainted dyads, whereas the respective loading of agreeableness continually decreased. However, the emergence of α depended on the dominating culture. Stability was identified in North American samples but was ill defined among European samples. The lack of invariance across culture has also been noted previously (cf. Jang et al., 2006) and makes the view of a universal two-factorial concept of personality seem premature. Rather, it seems prudent for future research to systematically examine the higher order structure of personality across different languages, countries and societies.

Limitations

Some caveats might limit the generalizability of these results to some degree. One limitation pertains to the methodological avenue adopted for this study. The meta-analysis relied on the reported correlations between the observed trait scores but had no access to the item-level data. Thus, instruments that do not have a pure simple factorial structure but include items that load on several trait factors could have created spurious correlations between the trait scores. Such factor blends can contribute to the emergence of artificial higher order factors of personality (Ashton, Lee, Goldberg, & deVries, 2009) because cross-loadings of selected items on two or more factors that are not accounted for result in spurious correlations between Big Five scores (Marsh et al., 2010). Thus, future studies are encouraged to replicate these findings with item-level data and to explicitly model latent constructs that can account for potential factor blends. In addition, the type of administered FFM instrument should be explicitly acknowledged. Although the g_p has been extracted from various scales (cf. Rusthon & Irwing, 2009), it seems to emerge less clearly in instruments using neutrally phrased items from which the socially desirable content has been removed (Backström et al., 2009). Another limitation of this study pertains to the modeling strategy of the latent factors.

Because the current study included only two informants (self and peer), it was necessary to constrain the loadings of the latent trait factors; thus, self- and observers ratings contributed equally to the latent trait variance. This limitation could be overcome in future studies by including more raters—for example, different types of observers (e.g., family members and friends). Furthermore, it is conceivable that the acquaintance effect might be confounded with developmental differences related to the age of the respondents: Short-term acquaintances were younger (*Mdn* = 20.4 years) than long-term acquaintances (*Mdn* = 33.5 years). As people mature, self-reports of personality generally become more differentiated across traits and result in lower correlations within the FFM (Soto, John, Gosling, & Potter, 2008). Stronger evidence could be gathered from matched samples with different lengths of acquaintance but comparable age structures. Moreover, this study examined only the quantitative aspect of acquaintance (i.e., its length) but neglected the qualitative component (i.e., the type of relationship; cf. Starzyk et al., 2006). It is possible that the degree of emotional attachment between raters results in attributions of more desirable characteristics to others than for targets with whom observers are not as strongly involved (Connelly & Ones, 2010). Finally, it should be acknowledged that most dyads in the aggregated primary studies were not randomized. Thus, a selection bias cannot be ruled out: it is conceivable that only pairs that agree on each others personality stay friends whereas dyads that disagree would cease their interactions and, thus, exhibit a shorter length of acquaintance. The combined effects of the quantitative and qualitative aspects of acquaintance should be examined more closely in future studies, for example, by using randomized roommate pairs (cf. Kurtz & Sherker, 2003).

Conclusion

As attractive as the idea of a general trait at the top of a personality hierarchy might seem, its empirical support based on the present meta-analysis is rather weak. Although the Big Five exhibited minor correlations between each other even when controlling for method-specific biases, they were rather small and, moreover, susceptible to the length of acquaintance between raters. A putative g_p could be extracted from ratings of dyads who had known each other for a comparably short period of time, but it gradually disappeared with increasing length of acquaintance. This sheds doubt on the idea that the g_p is a substantive trait that is more than a shared bias in self- and other ratings. More likely, the previously identified g_p across informants represents shared normative ratings that result from socially desirable responding in self-reports (Backström et al., 2009) and implicit personality theories in other reports (Beer & Watson, 2008a) rather than substantive aspects of an individual’s personality.

References

Articles included in the meta-analyses are listed in the supplement.

Alessandri, G., & Vecchione, M. (2012). The higher-order factors of the Big Five as predictors of job performance. *Personality and Individual Differences, 53*, 779-784. doi:10.1016/j.paid.2012.05.037

Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The Halo-Alpha-Beta model. *Journal of Personality and Social Psychology, 97*, 1142-1156. doi:10.1037/a0017159

Asthon, M. C., Lee, K., Goldberg, L. R., & de Vries, R. E. (2009). Higher order factors of personality: Do they exist? *Personality and Social Psychology Review, 13*, 79-91. doi:10.1177/1088868309338467

Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*, 533–548. doi:10.1037/a0016229

Backström, M. (2007). Higher-order factors in a five-factor personality inventory and its relation to social desirability. *European Journal of Psychological Assessment, 23*, 63-70. doi:10.1027/1015-5759.23.2.63

Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43*, 335-344. doi:10.1016/j.jrp.2008.12.013

Beer, A., & Watson, D. (2008a). Personality judgment at zero acquaintance: Agreement, assumed similarity, and implicit simplicity. *Journal of Personality Assessment, 90*, 250-260. doi:10.1080/00223890701884970

Beer, A., & Watson, D. (2008b). Asymmetry in judgments of personality: Others are less differentiated than the self. *Journal of Personality, 76*, 535-560. doi:10.1111/j.1467-6494.2008.00495.x

Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person perception accuracy: Another look at self-other agreement. *Personality and Social Psychology Bulletin, 20*, 367-378. doi:10.1177/0146167294204004

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self–other agreement in judgments of personality. *Journal of Personality and Social Psychology, 92*, 119-135. doi:10.1037/0022-3514.92.1.119

Block, J. (2010). The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry, 21*, 2-25. doi:10.1080/10478401003596626

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota symposia on child psychology* (Vol. 13, pp. 39-101). Hillsdale, NJ: Erlbaum.

Carney, D. A., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality, 41*, 1054-1072. doi:10.1016/j.jrp.2007.01.004

Carroll, J. B. (2002). The five-factor personality model: How complete and satisfactory is it? In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 97-126). Mahwah: Lawrence Erlbaum Associates Publishers.

Chang, L., Connelly, B. S., & Geeza, A. A. (2012). Separating method factors and higher order traits of the Big Five: A meta-analytic multitrait-multimethod approach. *Journal of Personality and Social Psychology*, *102*, 408-426. doi:10.1037/a0025559

Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods & Research*, *29*, 468-508. doi:10.1177/0049124101029004003

Cheung, S. F., & Chan, D. K.-S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology*, *89*, 780-91. doi:10.1037/0021-9010.89.5.780

Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: a two-stage approach. *Psychological Methods*, *10*, 40-64. doi:10.1037/1082-989X.10.1.40

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*, 1092-1122. doi:10.1037/a0021212

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*, 110-117. doi:10.1111/j.1468-2389.2007.00371.x

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R: Professional manual*. Odessa: Psychological Assessment Resources.

Cronbach, L. J. (1955). Process affecting scores on “understanding of others” and “assumed similarity”. *Psychological Bulletin, 52*, 177-193. doi:10.1037/h0044919

Danay, E., & Ziegler, M. (2011). Is there really a single factor of personality? A multirater approach to the apex of personality. *Journal of Research in Personality, 45,* 560-567. doi:10.1016/j.jrp.2011.07.003

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*, 1138-1151. doi:10.1037/0022-3514.91.6.1138

DeYoung, C., Hasher, L., Djikic, M., Criger, B., & Peterson, J. B. (2007). Morning people are stable people: Circadian rhythm and the higher-order factors of the Big Five. *Personality & Individual Differences, 43*, 267–276. doi:10.1016/j.paid.2006.11.030

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences*, *33*, 533-552. doi:10.1016/S0191-8869(01)00171-4

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal or Personality, 73*, 825-858. doi:10.1111/j.1467-6494.2005.00330.x

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41,* 417-440. doi:10.1146/annurev.ps.41.020190.002221

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*, 1246-1256. doi:10.1037/0022-3514.73.6.1246

Edwards, L. K., & Edwards, A. L. (1991). A principal-components analysis of the Minnesota Multiphasic Personality Inventory factor scales. *Journal of Personality and Social Psychology, 60*, 766-772. doi:10.1037/0022-3514.60.5.766

Erdle, S., Irwing, P., Rushton, J. P., & Park, J. (2010). The general factor of personality and its relation to self-esteem in 628,640 Internet respondents. *Personality and Individual Differences*, *48*, 343-346. doi:10.1016/j.paid.2009.09.004

Eysenck, H. J. (1947). *Dimensions of personality*. London: Routledge & Kegan-Paul.

Fletcher, G. J. O., & Kerr, P. S. G. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin, 136*, 627-658. doi:10.1037/a0019792

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethods matrix. *Psychological Bulletin, 56,* 81–105. doi:10.1037/h0046016

Furr, R. M. (2008). A framework for profile similarity: Integrating similarity, normativeness, and distinctiveness. *Journal of Personality, 76*, 1267-1316. doi:10.1111/j.1467-6494.2008.00521.x

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 2, pp. 141.165). Beverly Hills: Sage.

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: Tilburg University Press.

Hafdahl, A. R., Panter, A. T., Gramzow, R. H., Sedikides, C., & Insko, C. A. (2000). Free-response self-discrepancies across, among, and within FFM personality dimensions. *Journal of Personality, 68*, 111-151. doi:10.1111/1467-6494.00093

Hildebrandt, A., Sommer, W., Herzmann, G., & Wilhelm, O. (2010). Structural invariance and age-related performance differences in face cognition. *Psychology and Aging, 25*, 794-810. doi:10.1037/a0019774

Hirsh, J. B., DeYoung, C. G., & Peterson, J. B. (2009). Metatraits of the Big Five differentially predict engagement and restraint of Behavior. *Journal of Personality*, *77*, 1085-1102. doi:10.1111/j.1467-6494.2009.00575.x

Hopwood, C. J., Wright, A. G. C., & Donnellan, B. M. (2011). Evaluating the evidence for the general factor of personality across multiple inventories. *Journal of Research in Personality*, *45*, 468-478. doi:10.1016/j.jrp.2011.06.002

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis*. Thousand Oaks: Sage.

Jang, K. L., Livesley, W. J., Ando, J., Yamagata, S., Suzuki, A., Angleitner, A., … Spinath, F. (2006). Behavioral genetics of the higher-order factors of the Big Five. *Personality and Individual Differences*, *41*, 261–272. doi:10.1016/j.paid.2005.11.033

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York: Guilford Press.

Kurtz, J. E., & Sherker, J. L. (2003). Relationship quality, trait similarity, and self-other agreement on personality ratings in college roommates. *Journal of Personality, 71*, 21-48. doi:10.1111/1467-6494.t01-1-00005

Li, A., & Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self-deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment, 14,* 131-141. doi: 10.1111/j.1468-2389.2006.00339.x

Loehlin, J. C. (2011). Correlation between general factors for personality and cognitive skills in the National Merit twin sample. *Journal of Research in Personality*, *45*, 504-507. doi:10.1016/j.jrp.2011.06.011

Loehlin, J. C., & Horn, J. M. (2012). How general is the “General Factor of Personality”? Evidence from the Texas Adoption Project. *Journal of Research in Personality, 46,* 655-663. doi:10.1016/j.jrp.2012.07.004

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88,*139–157. doi:10.1037/0022-3514.88.1.139

Marsh, H. W., Lüdtke, O., Muthén, B. O., Asparouhov, T., Morin, A. J. S., & Trautwein, U. (2010). A new look at the Big-Five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471-491. doi:10.1037/a0019227

McCrae, R. R., Costa, P. T. Jr., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., Hřebíčková, M., & Urbánek, T. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality, 38*, 179-201. doi:10.1016/S0092-6566(03)00056-4

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*, 28-50. doi:10.1177/1088868310366253

McCrae, R. R., Yamagata, S., Jang, K. L., Riemann, R., Ando, J., Ono, Y., … Spinath, F. M. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *95*, 442–455. doi:10.1037/0022-3514.95.2.442

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah: Lawrence Erlbaum.

Mount, M., K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resources Management, 13*, 153-200.

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the Big Five personality traits and the big six vocational interest types. *Personnel Psychology, 58*, 447-478. doi:10.1111/j.1744-6570.2005.00468.x

Mowen, J. C., Park, S., & Zablah, A. (2007). Toward a theory of motivation and personality with application to word-of-mouth communications. *Journal of Business Research, 60*, 590-596. doi:10.1016/j.jbusres.2006.06.007

Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, *41*, 1213-1233. doi:10.1016/j.jrp.2007.02.003

Oh, I. S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, *96*, 762–773. doi:10.1037/a0021832

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63*, 816-824. doi:10.1037//0022-3514.63.5.816

Paulhus, D. L., Bruce, M. N., & Trapnell, P. D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality and Social Psychology Bulletin, 21*, 100-108. doi:10.1177/0146167295212001

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: More evidence for a “kernel of truth” in social perception. *Social Cognition, 24*, 607-640. doi:10.1521/soco.2006.24.5.607

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879-903. doi:10.1037/0021-9010.88.5.879

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*, 667-696. doi:10.1080/00273171.2012.715555

Revelle, W. (1979). Hierarchical cluster-analysis and the internal structure of tests. *Multivariate Behavioral Research, 14*, 57-74. doi:10.1207/s15327906mbr1401_4

Riemann, R., & Kandler, C. (2010), Construct validation using multitrait-multimethod-twin data: The case of a general factor of personality. *European Journal of Personality, 24*, 258-277. doi:10.1002/per.760

Robins, R. W, John, O. P., & Caspi, A. (1994). Major dimensions of personality in early adolescence: The Big Five and beyond. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 267-291). Hillsdale, NJ: Erlbaum.

Rushton, J. P., Bons, T. A., Ando, J., Hur, Y.-M., Irwing, P., Vernon, P. A., Petrides, K. V., et al. (2009). A general factor of personality from multitrait-multimethod data and crossnational twins. *Twin Research and Human Genetics*, *12*, 356-365. doi:10.1375/twin.12.4.356

Rushton, J. P., & Irwing, P. (2008). A General Factor of Personality (GFP) from two meta-analyses of the Big Five. *Personality and Individual Differences, 45*, 679-683. doi:10.1016/j.paid.2008.07.015

Rushton, J. P., & Irwing, P. (2009). A general factor of personality in 16 sets of the Big Five, the Guilford-Zimmerman Temperament Survey, California Psychological Inventory, and Temperament and Character Inventory. *Personality and Individual Differences, 47*, 558-564. doi:10.1016/j.paid.2009.05.009

Rushton, J. P., & Irwing, P. (2011). The general factor of personality: Normal and abnormal. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell Handbook of Individual Differences* (pp. 132-161). London: Blackwell.

Salgado, S. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, *76*, 323-346. doi:10.1348/096317903769647201

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online, 8*, 23-74.

Schimmack, U. (2010). What multi-method data tell us about construct validity. *European Journal of Personality, 24,* 241–257. doi:10.1002/per.771

Schmid, J. Jr., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 83-90. doi:10.1007/BF02289209

Schmidt, F. L., Oh, I. S., & Hayes, T. L. (2009). Fixed-versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology, 62*, 97-128. doi:10.1348/000711007X255327

Schmit, M. J., & Ryan, A.M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology, 78*, 966-974. doi:10.1037/0021-9010.78.6.966

Schneider, L., Schimmack, U., Petrican, R., & Walker, S. (2010). Acquaintanceship length as a moderator of self-informant agreement in life satisfaction ratings. *Journal of Research in Personality, 44*, 146-150. doi:10.1016/j.jrp.2009.11.004

Simms, L. J., Zelazny, K., Yam, W. H., Gros, D. F. (2010). Self-informant agreement for personality and evaluative person descriptors: Comparing methods for creating informant measures. *Journal of Personality, 24*, 207-221. doi:10.1002/per.763

Şimşek, Ö. F., Koydemir, S., & Schütz, A. (2012). A multigroup multitrait-multimethod study in two countries supports the validity of a two-factor higher order model of personality. *Journal of Research in Personality, 46*, 442-449. doi:10.1016/j.jrp.2012.04.005

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718-737. doi:10.1037/0022-3514.94.4.718

Starzyk, K. B., Holden, R. R., Fabrigar, L. R., & MacDonald, T. K.(2006). The personal acquaintance measure: A tool for appraising one's acquaintance with any person. *Journal of Personality and Social Psychology, 90,* 833-847. doi:10.1037/0022-3514.90.5.833

Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, *44*, 315-327. doi:10.1016/j.jrp.2010.03.003

Van der Linden, D., Scholte, R. H. J., Cillessen, A. H. N., te Nijenhuis, J., & Segers, E. (2010). Classroom ratings of likeability and popularity are related to the Big Five and the general factor of personality. *Journal of Research in Personality, 44*, 669–672. doi:10.1016/j.jrp.2010.08.007

Van der Linden, D., te Nijenhuis, J., Cremers, M., & Van de Ven, C. (2011). General factors of personality in six datasets and a criterion-related validity study at the Netherlands Armed Forces. *International Journal of Selection and Assessment*, *19*, 157-169. doi:10.1111/j.1468-2389.2011.00543.x

Vazire, S., & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science, 20*, 104-108. doi:10.1177/0963721411402478

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*, 1202–1216. doi:10.1037/a0013314

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*, 261-293. doi:10.3102/10769986030003261

Viechtbauer, W., & Cheung, W. (2010). Outlier and influencer diagnostics for meta-analysis. *Research Synthesis Methods, 1*, 110-125. doi:10.1002/jrsm.11

Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology, 48*, 865-885. doi:10.1111/j.1744-6570.1995.tb01784.x

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The Role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, *78*, 546-558. doi:10.1037/0022-3514.78.3.546

Wood, D., Gosling, S. D., & Potter, J. (2007). Normality evaluations and their relation to personality traits and well-being. *Journal of Personality and Social Psychology, 93,* 861-879. doi:10.1037/0022-3514.93.5.861

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach ’s α, Revelle ’s β, and McDonald ’s ω_H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*, 123-133. doi:10.1007/s11336-003-0974-7

Table 1

Aggregated Multi-Informant Correlations for Big Five Traits

			Self-report					Peer report							
			C	A	N	O	E	C	A	N	O	E			
Uncorrected mean correlations	Self-report $\lvert\bar{r}\rvert = .17$	C		.29* (.14)	-.29* (.10)	.02 (.07)	.15* (.00)	**.52* (.13)**	.02 (.07)	-.08 (.07)	-.08 (.10)	.01* (.00)	C	Self report $\lvert\bar{r}\rvert = .21$	Corrected mean correlations
		A	.22 (.13)		-.31* (.11)	.15 (.12)	.22* (.08)	.10 (.06)	**.46* (.17)**	-.07 (.08)	.02* (.01)	.07* (.03)	A		
		N	-.25* (.11)	-.24* (.11)		-.12* (.05)	-.32* (.03)	-.12 (.07)	-.10* (.05)	**.43* (.18)**	.01* (.00)	-.13 (.14)	N		
		O	.03 (.09)	.12 (.12)	-.11 (.09)		.24 (.12)	-.07* (.00)	-.02 (.05)	.02 (.08)	**.50* (.20)**	.09* (.03)	O		
		E	.11 (.06)	.17 (.10)	-.25* (.07)	.20 (.12)		-.02* (.00)	.04 (.11)	-.12 (.17)	.06 (.05)	**.59* (.12)**	E		
	Peer report $\lvert\bar{r}\rvert = .21$	C	**.43* (.13)**	.08 (.09)	-.10 (.11)	-.06 (.06)	.00 (.10)		.34* (.11)	-.29* (.11)	.22 (.14)	.09 (.15)	C	Peer report $\lvert\bar{r}\rvert = .26$	
		A	.02 (.10)	**.36* (.16)**	-.08 (.11)	.00 (.10)	.04 (.14)	.27* (.11)		-.42* (.15)	.32* (.08)	.20* (.07)	A		
		N	-.07 (.10)	-.06 (.11)	**.35* (.17)**	.01 (.11)	-.11 (.16)	-.24* (.12)	-.33* (.14)		-.18 (.09)	-.29 (.14)	N		
		O	-.08 (.12)	.03 (.07)	.00 (.08)	**.40* (.18)**	.05 (.08)	.19 (.14)	.26* (.09)	-.16 (.10)		.27* (.11)	O		
		E	.02 (.09)	.07 (.08)	-.11 (.14)	.08 (.10)	**.51* (.12)**	.07 (.15)	.16 (.09)	-.24 (.14)	.23 (.12)		E		

Note. C = Conscientiousness, A = Agreeableness, N = Neuroticism, O = Openness, E = Extraversion; Uncorrected (below diagonal) and corrected (above diagonal) mean correlations with standard deviations in parentheses. Convergent correlations (same trait, different informant) are in boldface. Corrections include adjustments for random and transient error.

* $p < .05$ based on the 95% credibility interval.

Table 2

Self-Other Agreement at Short- and Long-Term Acquaintance

	Short-term acquaintance		Long-term acquaintance		
	r	*SE*	*r*	*SE*	Δr
Conscientiousness	.44*	.03	.58*	.02	.13*
Agreeableness	.38*	.03	.53*	.03	.15*
Neuroticism	.33*	.04	.50*	.02	.17*
Openness	.36*	.03	.59*	.03	.22*
Extraversion	.55*	.03	.62*	.02	.08*

Note. *SE* = Standard error; Δr = Difference in correlations (long-term minus short-term).
* $p < .05$.

Table 3

Latent Multi-Informant Correlations at Short- and Long-Term Acquaintance

	Short-term acquaintance		Long-term acquaintance		
	r	*SE*	*r*	*SE*	Δ*r*
C ↔ A	.13*	.05	-.04	.03	-.17*
C ↔ N-	.19*	.04	.07*	.02	-.11*
C ↔ O	.19*	.04	.10*	.02	-.09*
C ↔ E	.26*	.05	.16*	.03	-.10*
A ↔ N-	.20*	.04	.12*	.02	-.08*
A ↔ O	-.20*	.05	-.12*	.02	.09*
A ↔ E	.04	.05	-.06*	.03	-.09*
N- ↔ O	.06	.04	-.02	.02	-.08*
N- ↔ E	.30*	.04	.22*	.02	-.07*
O ↔ E	.11*	.05	.23*	.02	.12*

Note. *SE* = Standard error; Δ*r* = Difference in correlations (long-term minus short-term); C = Conscientiousness, A = Agreeableness, N- = Neuroticism (reverse scored), O = Openness, E = Extraversion.

* $p < .05$.

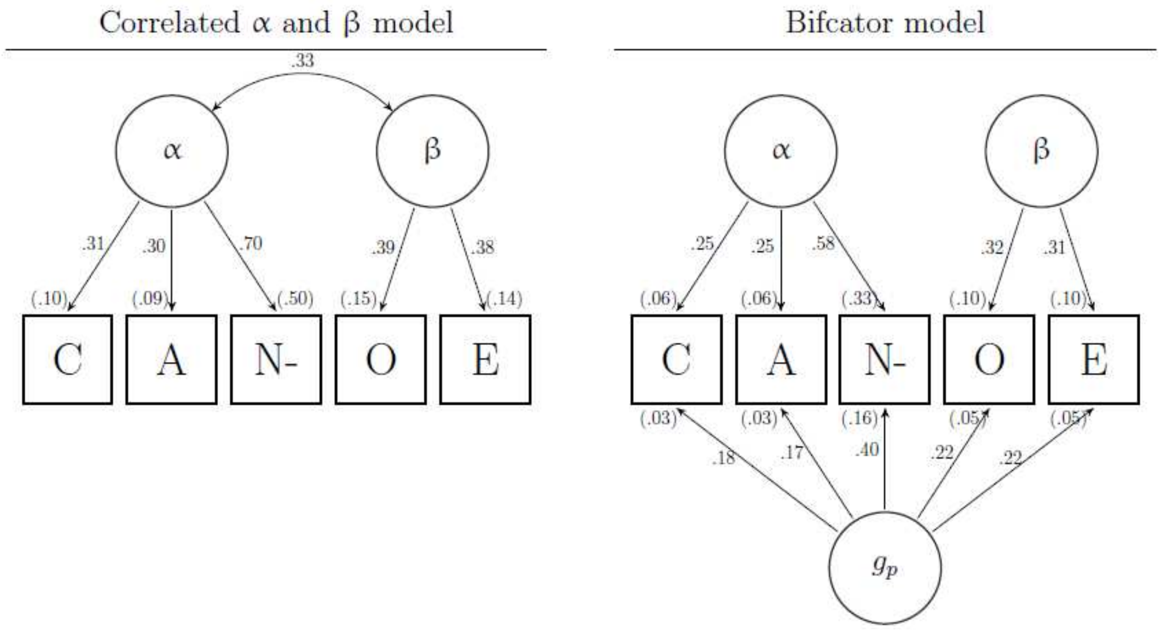

Figure 1. Standardized factor loadings and variance explained (in parentheses) of general and specific factors in latent multi-informant ratings. α = Stability, β = Plasticity, g_p = General factor of personality, C = Conscientiousness, A = Agreeableness, N- = Neuroticism (reverse scored), O = Openness, E = Extraversion. Measurement models are not presented.

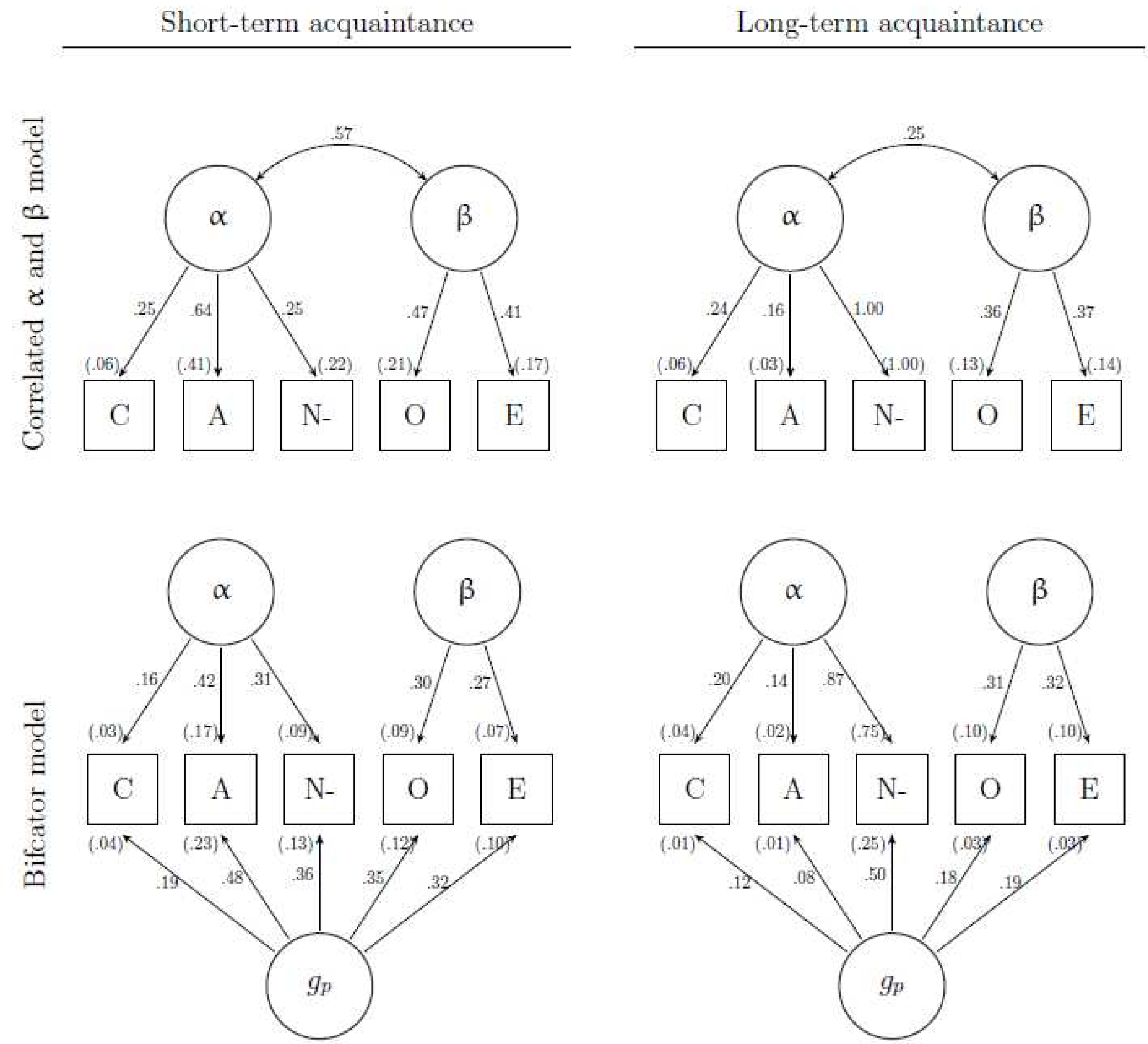

Figure 2. Standardized factor loadings and variance explained (in parentheses) of higher order factors in latent multi-informant ratings at short- (*Mdn* = 6 months) and long-term acquaintance (*Mdn* = 14 years). α = Stability, β = Plasticity, g_p = General factor of personality, C = Conscientiousness, A = Agreeableness, N- = Neuroticism (reverse scored), O = Openness, E = Extraversion. Measurement models are not presented.

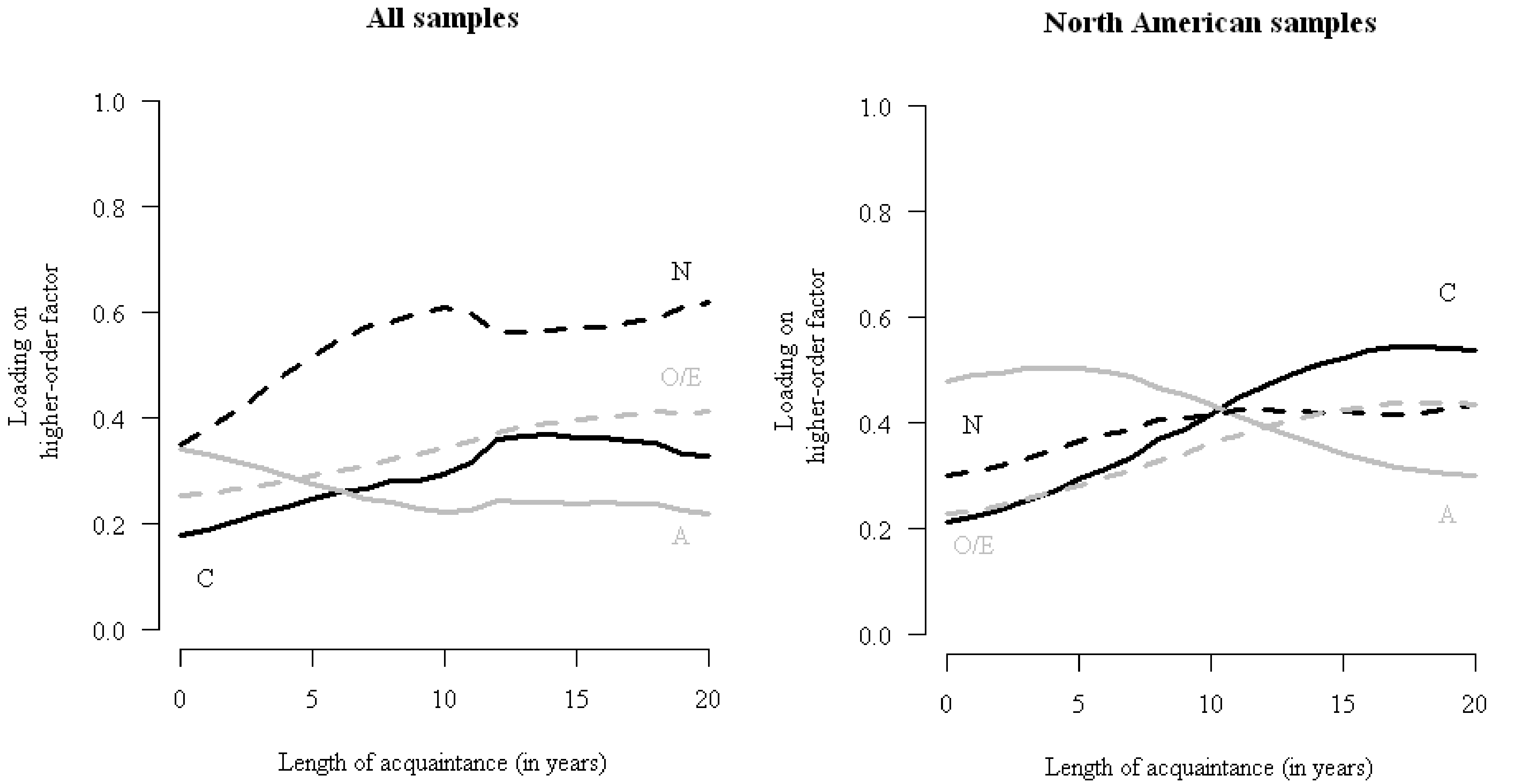

Figure 3. Factor loading on higher order factors of personality in multi-informant ratings from 0 to 20 years of acquaintance. C = Conscientiousness, A = Agreeableness, N- = Neuroticism (reverse scored), O / E= Openness / Extraversion.

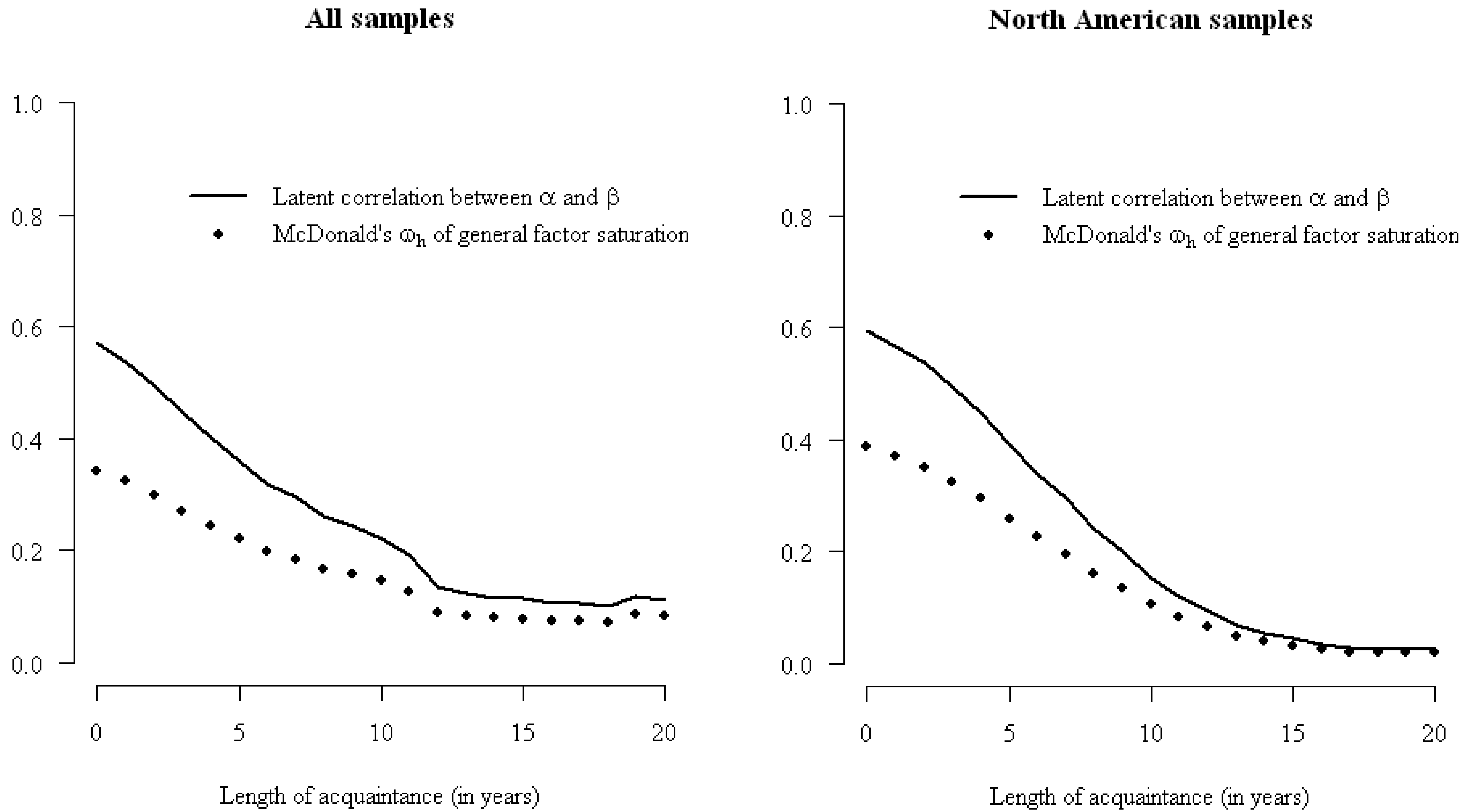

Figure 4. Latent correlations of higher order factors of personality in multi-informant ratings from 0 to 20 years of acquaintance.

Online Supplement for

"The Elusive General Factor of Personality: The Acquaintance Effect"

Timo Gnambs

Osnabrück University

Calculation of locally weighted averages

Gradients of focal parameters, for example the correlation between the two higher order factors α and β, were estimated by weighing each individual effect size and recalucating the meta-analyses and subsequent latent variable models on the weighted correlations. Adopting a procedure similar to nonparametric regression (Li & Racine, 2007) or local structural equation modeling (Hildebrandt, Sommer, Herzmann, & Wilhelm, 2010), weights were determined using a Gaussian kernel function (Silverman, 1986) based on the average length of acquaintance between informants reported for each sample:

Let r_i denote the vector of correlations for one of the 45 meta-analyses presented above and x_i the average length of acquaintance reported for the corresponding samples. The meta-analytically derived true score correlation for the r_i at a focal point x_{focal}, for example at one year of acquaintance, could be estimated by selecting a subsample of r_i where $x_i = x_{focal}$. However, this would require an appropriately large sample at the defined x_{focal}. Instead, locally weighted estimations determine weights w_i for each r_i that approach 1 at $x_i \approx x_{focal}$ and asymptotically near 0 for $x_i < x_{focal}$ and $x_i > x_{focal}$, and then calculate the true score correlation on all r_i weighted by w_i. In the latter case, all r_i contribute to the estimated true score correlation at a defined x_{focal} depending on the distance of the samples' x_i from x_{focal}. The weights w_i were determined using a kernel function following a standard normal distribution (Silverman, 1986):

[1] $$w_i = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z_i^2}$$

The scaled distance between each x_i and the defined x_{focal} is given as

[2] $$z_i = (x_i - x_{focal}) / h$$

where the smoothing parameter h is optimally approximated by

[3] $$h = \left(\frac{4 \cdot SD_x^5}{3 \cdot n} \right)^{\frac{1}{5}}$$

By (a) estimating w_i at each $x_\mathrm{focal} \in \{0 \text{ to } 20 \text{ years}\}$, (b) recalculating the respective meta-analyses using r_i weighted by w_i, and (c) estimating the latent variable models presented above on these true score correlations, gradients for various parameters of interests can be derived; for example, the latent correlation between the two higher order factors of personality, α and β, across different lengths of acquaintance.

Hildebrandt, A., Sommer, W., Herzmann, G., & Wilhelm, O. (2010). Structural invariance and age-related performance differences in face cognition. *Psychology and Aging, 25*, 794-810. doi:10.1037/a0019774

Li, Q., & Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton: University Press.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman & Hall.

Supplemental Tables

Table S1

Self-Other Agreement at Short- and Long-Term Acquaintance in North American Samples

	Short-term acquaintance		Long-term acquaintance		
	r	*SE*	*r*	*SE*	Δ*r*
Conscientiousness	.48*	.02	.60*	.02	.12*
Agreeableness	.36*	.03	.58*	.04	.22*
Neuroticism	.32*	.04	.54*	.03	.22*
Openness	.34*	.04	.64*	.03	.30*
Extraversion	.56*	.03	.64*	.03	.07*

Note. *SE* = Standard error; Δ*r* = Difference in correlations (long-term minus short-term).

* $p < .05$.

Table S2

Latent Multi-Informant Correlations at Short- and Long-Term Acquaintance in North American Samples

	Short-term acquaintance		Long-term acquaintance		
	r	*SE*	*r*	*SE*	Δ*r*
C ↔ A	.10	.06	-.03	.03	-.13*
C ↔ N-	.26*	.05	.12*	.03	-.14*
C ↔ O	.21*	.05	.16*	.03	-.05
C ↔ E	.33*	.05	.19*	.03	-.14*
A ↔ N-	.20*	.04	.11*	.03	-.09*
A ↔ O	-.16*	.06	-.08*	.02	.09*
A ↔ E	.07	.06	-.11*	.03	-.18*
N- ↔ O	.14	.04	.01	.03	-.12*
N- ↔ E	.32*	.05	.20*	.03	-.11*
O ↔ E	.12*	.05	.25*	.03	.13*

Note. *SE* = Standard error; Δ*r* = Difference in correlations (long-term minus short-term); C = Conscientiousness, A = Agreeableness, N- = Neuroticism (reverse scored), O = Openness, E = Extraversion.

* $p < .05$.

Articles included in the meta-analysis

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*, 340-345. doi:10.1080/00223890902935878

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21*, 372-374. doi:10.1177/0956797609360756

Beer, A., & Watson, D. (2008a). Asymmetry in judgments of personality: Others are less differentiated than the self. *Journal of Personality, 76*, 535-560. doi:10.1111/j.1467-6494.2008.00495.x

Beer, A., & Watson, D. (2008b). Personality judgment at zero acquaintance: Agreement, assumed similarity, and implicit simplicity. *Journal of Personality Assessment, 90,* 250-260. doi:10.1080/00223890701884970

Biesanz, J. C., & West, S. G. (2004). Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality, 72,* 845-876. doi:10.1111/j.0022-3506.2004.00282.x

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and celf–other agreement in judgments of personality. *Journal of Personality and Social Psychology, 92,* 119-135. doi:10.1037/0022-3514.92.1.119

Borkenau, P., & Liebler, A. (1993a). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology, 65,* 546-553. doi:10.1037/0022-3514.65.3.546

Borkenau, P., & Liebler, A. (1993b). Consensus and self-other agreement for trait inferences from minimal information. *Journal of Personality, 61,* 477-496. doi:10.1111/j.1467-6494.1993.tb00779.x

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology, 86*, 599-614. doi:10.1037/0022-3514.86.4.599

Connelly, B. S., & Hülsheger, U. R. (2011). A narrower scope or a clearer lens for personality? Examining sources of observers' advantages over self-reports for predicting performance. *Journal of Personality, 80*, 603-631. doi:10.1111/j.1467-6494.2011.00744.x

Durbin, C. E., Schalet, B. D., Hayden, E. P., Simpson, J., & Jordan, P. L. (2009). Hypomanic personality traits: A multi-method exploration of their association with normal and abnormal dimensions of personality. *Journal of Research in Personality, 43*, 898-905. doi:10.1016/j.jrp.2009.04.010

Edwards, B. D., & Woehr, D. J. (2007). An examination and evaluation of frequency-based personality measurement. *Personality and Individual Differences, 43,* 803-814. doi:10.1016/j.paid.2007.02.005

Foltz, C., Morse, J. Q., Calvo, N., & Barber, J. P. (1997). Self- and observer ratings on the NEO-FFI in couples: Initial evidence of the psychometric properties of an observer form. *Assessment, 4*, 287-295.

García, L. F., Aluja, A., García, Ó., & Colom, R. (2007). Do parents and children know each other? A study about agreement on personality within families. *Piscothema, 19*, 120-123.

Holland, A. S., & Roisman, G. I. (2008). Big Five personality traits and relationship quality: Self-reported, observational, and physiological evidence. *Journal of Social and Personal Relationships, 25*, 811-829. doi:10.1177/0265407508096697

Hong, R. Y., Koh, S., & Paunonen, S. V. (2012). Supernumerary personality traits beyond the Big Five: Predicting materialism and unethical behavior. *Personality and Individual Differences, 53,* 710-715. doi:10.1016/j.paid.2012.05.030

Humrichouse, J. J. (2010). *The hierarchical structure of emotional expressivity: Scale development and nomological implications* (Doctoral thesis, University of Iowa). Retrieved from http://ir.uiowa.edu/etd/519/

Jackson, B., Dimmock, J. A., Gucciardi, D. F., & Grove, J. R. (2010). Relationship commitment in athletic dyads: Actor and partner effects for Big Five self- and other-ratings. *Journal of Research in Personality, 44,* 641-648. doi:10.1016/j.jrp.2010.08.004

Joseph, D. L., & Newman, D. A. (2010). Discriminant validity of self-reported emotional intelligence: A multitrait-multisource study. *Educational and Psychological Measurement, 70,* 672-694. doi:10.1177/0013164409355700

Kamdar, D., & van Dyne, L. (2007). The joint effects of personality and workplace social exchange relationships in predicting task performance and citizenship performance. *Journal of Applied Psychology, 92,* 1286-1298. doi:10.1037/0021-9010.92.5.1286

Kurtz, J. E., Tarquinie, S. J., Lobst, E. A. (2008). Socially desirable responding in personality assessment: Still more substance than style. *Personality and Individual Differences, 45*, 22-27. doi:10.1016/j.paid.2008.02.012

Laidra, K., Alik, J., Harro, M., Merenäkk, L., & Harro, J. (2006). Agreement among adolescents, parents, and teachers on adolescent personality. *Assessment, 13,* 187-196. doi:10.1177/1073191106287125

Lee, K., & Ashton, M. C. (2006). Further assessment of the HEXACO Personality Inventory: Two new facet scales and an observer report form. *Psychological Assessment, 18,* 182-191. doi:10.1037/1040-3590.18.2.182

Lee, K., Ashton, M. C., Pozzebon, J. A., Visser, B. A., Bourdage, J. S., & Ogunfowora, B. (2009). Similarity and assumed similarity in personality reports of well-acquainted persons. *Journal of Personality and Social Psychology, 96,* 460-472. doi:10.1037/a0014059

Lee, K., Ashton, M. C., Morris, D. L., Cordery, J., & Dunlop, P. D. (2008). Predicting integrity with the HEXACO personality model: Use of self- and observer reports. *Journal of Occupational and Organizational Psychology, 81,* 147-167. doi:10.1348/096317907X195175

Lonnqvist, J.-E., Paunonen, S., Tuulio-Henriksson, A., Lonnqvist, J., & Verkasalo, M. (2007). Substance and style in socially desirable responding. *Journal of Personality, 75*, 291-322. doi:10.1111/j.1467-6494.2006.00440.x

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the Five-Factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81-90. doi:10.1037/0022-3514.52.1.81

McCrae, R. R., Stone, S. V., Fagan, P. J., & Costa Jr., P. T. (1998). Identifying causes of disagreement between self-reports and spouse ratings of personality. *Journal of Personality, 66*, 285-313. doi:10.1111/1467-6494.00013

Mlacic, B., & Ostendorf, F. (2005). Taxonomy and structure of Croatian personality-descriptive adjectives. *European Journal of Personality, 19,* 117-152. doi:10.1002/per.539

Muck, P. M., Hell, B., & Gosling, S. D. (2007). Construct validation of a short five-factor model instrument: A self-peer study on the German adaptation of the Ten-Item Personality Inventory (TIPI-G). *European Journal of Psychological Assessment, 23,* 166-175. doi:10.1027/1015-5759.23.3.166

O'Rourke, N., Neufeld, E., Claxton, A., & Smith, J. A., Z. (2010). Knowing me – knowing you: Reported personality and trait discrepancies as predictors of marital idealization between long-wed spouses. *Psychology and Aging, 25*, 412-421. doi:10.1037/a0017873

Park, B., Kraus, S., & Ryan, C. S. (1997). Longitudinal changes in consensus as a function of acquaintance and agreement in liking. *Journal of Personality and Social Psychology, 72*, 604-616. doi:10.1037/0022-3514.72.3.604

Parker, W., & Stumpf, H. (1998). A validation of the Five-Factor Model of personality in academically talented youth across observers and instruments. *Personality and Individual Differences, 25*, 1005-1025. doi:10.1016/S0191-8869(98)00016-6

Paulhus, D. L., & Bruce, M. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63*, 816-824. doi:10.1037/0022-3514.63.5.816

Piedmont, R. L. (1994). Validation of the NEO PI-R observer form for college students: Toward a paradigm for studying personality development. *Assessment, 1*, 259-268. doi:10.1177/107319119400100306

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality, 65*, 449-475. doi:10.1111/j.1467-6494.1997.tb00324.x

Romero, E., Villar, P., Gómez-Fraguela, J. A., & López-Romero, L. (2012). Measuring personality traits with ultra-short scales: A study of the Ten Item Personality Inventory (TIPI) in a Spanish sample. *Personality and Individual Differences, 53*, 289-293. doi:10.1016/j.paid.2012.03.035

Rouse, S. V., & Haas, H. A. (2003). Exploring the accuracies and inaccuracies of personality perception following Internet-mediated communication. *Journal of Research in Personality, 37*, 446-467. doi:S0092-6566(03)00025-4

Saroglou, V., & Fiasse, L. (2003). Birth order, personality, and religion: A study among young adults from a three-sibling family. *Personality and Individual Differences, 35*, 19-29. doi:10.1016/S0191-8869(02)00137-X

Simms, L. J., Zelazny, K., Yam, W. H., & Gros, D. F. (2010). Self-informant agreement for personality and evaluative person descriptors: Comparing methods for creating informant measures. *European Journal of Personality, 24*, 207-221. doi:10.1002/per.763

Small, E. E., & Diefendorff, J. M. (2006). The impact of contextual self-ratings and observer ratings of personality on the personality-performance relationship. *Journal of Applied Social Psychology, 36*, 297-320. doi:10.1111/j.0021-9029.2006.00009.x

Spinath, F. M. (2000). *Validität von Fremdbeurteilungen: Einflussfaktoren auf die Konvergenz von Selbst- und Fremdbeurteilungen in Persönlichkeitsschätzungen* [Validity of other-ratings: Factors influencing the convergence of self- and other-ratings of personality]. Lengrich: Pabst.

Spirrison, C. L., & Choi, S. (1998). Psychometric properties of a Korean version of the revised NEO-Personality Inventory. *Psychological Reports, 83*, 263-274. doi:10.2466/pr0.1998.83.1.263

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120-128. doi:10.1037/0022-3514.57.1.120

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*, 546-558. doi:10.1037/0022-3514.78.3.546

Westerlund, J., & Hansen, N. (2009). Validation of the Swedish version of the NEO-PI-R: Correlations between self-reports and college peer ratings. *Psychological Reports, 105,* 815-824. doi:10.2466/PR0.105.3.815-824

White, G. A. (2010). *Implications of relationship social comparison tendencies among dating and married individuals* (Doctoral thesis, University of Iowa, US). Retrieved from http://ir.uiowa.edu/etd/905/