Exploring the Etiology and Ontology of the Big Five Personality Traits

Multivariate Genetic Analysis of a Norwegian Sample

Joakim Coleman Ebeltoft and Jørgen Vorset Heiberg

Thesis submitted at the Department of Psychology, University of Oslo

Spring 2021

Summary

Authors: Joakim Coleman Ebeltoft and Jørgen Vorset Heiberg.
Title: Exploring the Etiology and Ontology of the Big Five Personality Traits. Multivariate Genetic Analysis of a Norwegian sample.
Supervisor: Nikolai Olavi Czajkowski.
Co-supervisors: Olav Vassend and Espen Røysamb.
Data: Data from 1272 twins recruited from the *Norwegian Twin Registry* were made available to us by our supervisors. The data were collected as part of a larger project researching health complaints, personality, and quality of life. The current thesis has no other affiliations to this project.

The general objective of this thesis is to apply modern multivariate methods in a Norwegian twin sample in order to sharpen and nuance crucial aspects of how we understand the reality and fundamental properties of the Big Five personality traits (i.e., the *ontology*) and their origin and development (i.e., the *etiology*).

To achieve the general objective, our first aim was to assess the Big Five facets, which were used as building blocks in the models applied in our analysis. This appraisal was performed in two steps. The first step was to examine the phenotypic facet dimensionality in the current sample. A five-factor structure emerged in our sample. However, significant alterations of two of the proposed main domains were evident, augmenting questions about the alleged universality of the Big Five traits. The second step was to estimate the heritability and its relation to construct unity. The moderate correlation between the wide-ranging heritability and Cronbach's alpha estimates, together with the apparent cross-loadings, raise essential ontological questions regarding the applied building blocks of the Big Five.

The second aim was to contribute to the knowledge of the traits' etiology by investigating the sources of variation underpinning the expression of personality. To address this aim, additive genetic and non-shared environmental facet correlation matrices were extracted from a Cholesky twin design model, and principal component analysis was applied to estimate factors. Five factors were extracted in the genetic correlation matrix, which greatly resembled the phenotypic rotation. Four factors emerged in the environmental matrix, which also resembled the phenotypic rotation with the exception of the fusion of two factors.

The third aim targeted the ontology of the Big Five traits, through scrutiny of the fundamental interpretation and application of the five-factor model. Ever since Allport (1931)

postulated that “a trait has more than nominal existence” (p. 368), the debate regarding the nature of personality traits has, to a certain extent, been buried by an implicit or explicit acceptance of Allport’s position. Although buried, the unsettled assumption has haunted the research field, and in this thesis, we dig up this nearly one-hundred-year-old statement by comparing two profound theoretical perspectives, which can be regarded as contrasting ontological positions: the *realist interpretation*, which considers the Big Five dimensions to be *veridical entities* that coincide with reality and emerge from a biological basis, ultimately anchored in genes, and the *constructivist interpretation*, which assumes that personality traits identified through factor analysis mainly reflect semantic clusters in the language. To assess these interpretations, common and independent pathways models were compared to test the five-factor model’s ability to mediate genetic and environmental contributions. The independent pathways model fits the data comparatively better, indicating that the five factors do not mediate genetic and environmental contributions to personality trait constructs. The model that fit the quantitative genetic data best was an alternative local etiological independent pathways model.

The results reaffirmed the ambiguous universality and equivocal facets of the Big Five, which question the origin, development, structure, and nature of the traits as proposed in the five-factor theory. The etiological exploration of genetic and environmental components indicates that both endogenous and systematic exogenous influence structure the Big Five traits. Ultimately, the results from the pointed operationalization of fundamental ontological positions are unsupportive of an interpretation of the Big Five traits as causal explanations for aggregated thoughts, emotions, and behavior.

Table of Contents

1. Introduction

1.1. A Recapitulation of Trait Psychology and the Five-Factor Model

Personality trait research revolves around identifying quantifiable dimensions that describe variation in human tendencies of feelings, thoughts, and behavior across situations and time. Several models seek to grasp the variation in these tendencies, such as Eysenck and Eysenck's (1991) three dimensional model and Hathaway and McKinley's (1951) Minnesota Multiphasic Personality Inventory. The *five-factor model* (FFM), often referred to as the *Big Five*, has been favored as a measure of personality in many bodies of literature and is the dominating taxonomy of normal personality variation (Widiger, 2017). The FFM consists of five personality dimensions: *neuroticism, extraversion, openness to experience, agreeableness,* and *conscientiousness,* often abbreviated as N, E, O, A, and C. The body of supportive literature is massive and growing. Nevertheless, the understanding of the Big Five's etiology and ontology remains ambiguous and debated. In this thesis, we seek to illuminate the understanding through methodological novelty and epistemological discussions.

1.1.2. It *Can* be Described with Words—the Psycholexical Approach

To be familiar with its epistemological foundation is essential when exploring a model. Although all three aforementioned models were developed with a combination of theoretical and empirical approaches, there are distinct differences in their epistemological basis. Eysenck (1963) stressed the importance of biological processes in the development of personality, while the Minnesota Multiphasic Personality Inventory was anchored in clinical psychology and psychiatry (Buchanan, 1994). The FFM, on the other hand, was largely founded on the empirical grounds of the psycholexical approach.

The psycholexical approach consists of three steps: i) extracting personality-descriptive words from a dictionary, ii) semantically reducing the number of descriptions, and iii) identifying overarching factors by applying factor analysis on a dataset of trait ratings. Proponents of the psycholexical approach assume that as hypersocial beings, humans develop language containing a rich variety of descriptions of personality differences in order to navigate the social world. This assumption is formulated as the psycholexical hypothesis: "(...) the more important such a difference, the more likely is it to become expressed as a single word" (John et al., 1988, p. 144). The notable psycholexical contributors Allport and Odbert (1936) reduced 18000 person-describing adjectives and nouns from the English dictionary to 4500 by excluding synonyms and non-neutral terms. From the 4500 words, the pioneer Cattell (1944) identified 171 dimensional traits. These traits serve as a foundation for many dimensional personality models, amongst them Cattell's (1966) own 16 personality factors and Norman's (1964) five factors, which among other solutions were an essential part of Costa and McCrae's (2008) development of the FFM.

Modern psychometrics relies heavily on the factor analytic method and an assumption that a latent and common variable explains much of the variation in the items. This applies to the field of personality trait research and the development of the FFM as well. By correlating the item responses, identification of proportions of shared variance (communality) is possible. To settle on a number of factors that maintain both parsimony and explanatory power is the main challenge in model and instrument development. As pragmatically pointed out by Costa and McCrae (1992, p. 180): "Surely every personality test would have omissions (...) How many traits should an inventory measure?"

1.1.3. Five Words are Enough—the Five-Factor Model

In the decades following the late 1940s, several research groups applied the psycholexical approach and found a wide variety of factor solutions. During the 1980s, a substantial proportion of researchers converged on the five-factor structure, although there were disputes regarding the naming and contents of the fifth factor (*openness to experience/intellect/culture*) (Goldberg, 1993). Costa and McCrae (2008) were important contributors in this process and showed the overlap between their FFM and many of the other factor structures, which strengthened the notion that the five factors were the most versatile. A central contribution was the development of the Revised NEO Personality Inventory (NEO PI-R) to measure the five factors (Costa & McCrae, 1992).

Several researchers also opposed the FFM for various reasons. For example, Paunonen and Jackson (2000) argued for additional factors in the English language. They reported that nine personality dimensions were left out of the Big Five model. Block (1995) was also concerned about the items that were left out of the questionnaires because of a lack of communality with other items. The importance of a factor or item is measured by the communality, which makes the questionnaires vulnerable to exclude important traits that do not share the largest proportions of communality. On the other hand, sharing a large proportion of communality variables of little real-world relevance could be interpreted as important.

Despite the critique, the NEO PI-R is widely applied in numerous contexts, such as research and personnel selection (Widiger, 2017). The current thesis is no exception. The NEO PI-R measures the five N, E, O, A, and C factors and six subdomains referred to as facets.

Table 1

Domains and facets of the FFM

Neuroticism (N)	Extraversion (E)	Openness to Experience (O)	Agreeableness (A)	Conscientiousness (C)
N1 Anxiety	E1 Warmth	O1 Fantasy	A1 Trust	C1 Competence
N2 Angry Hostility	E2 Gregariousness	O2 Aesthetics	A2 Straightforwardness	C2 Order
N3 Depression	E3 Assertiveness	O3 Feelings	A3 Altruism	C3 Dutifulness
N4 Self-Consciousness	E4 Activity	O4 Actions	A4 Compliance	C4 Achievement Striving
N5 Impulsiveness	E5 Excitement-Seeking	O5 Ideas	A5 Modesty	C5 Self-Discipline
N6 Vulnerability	E6 Positive Emotion	O6 Values	A6 Tender-Mindedness	C6 Deliberation

Note. Costa and McCrae (1992).

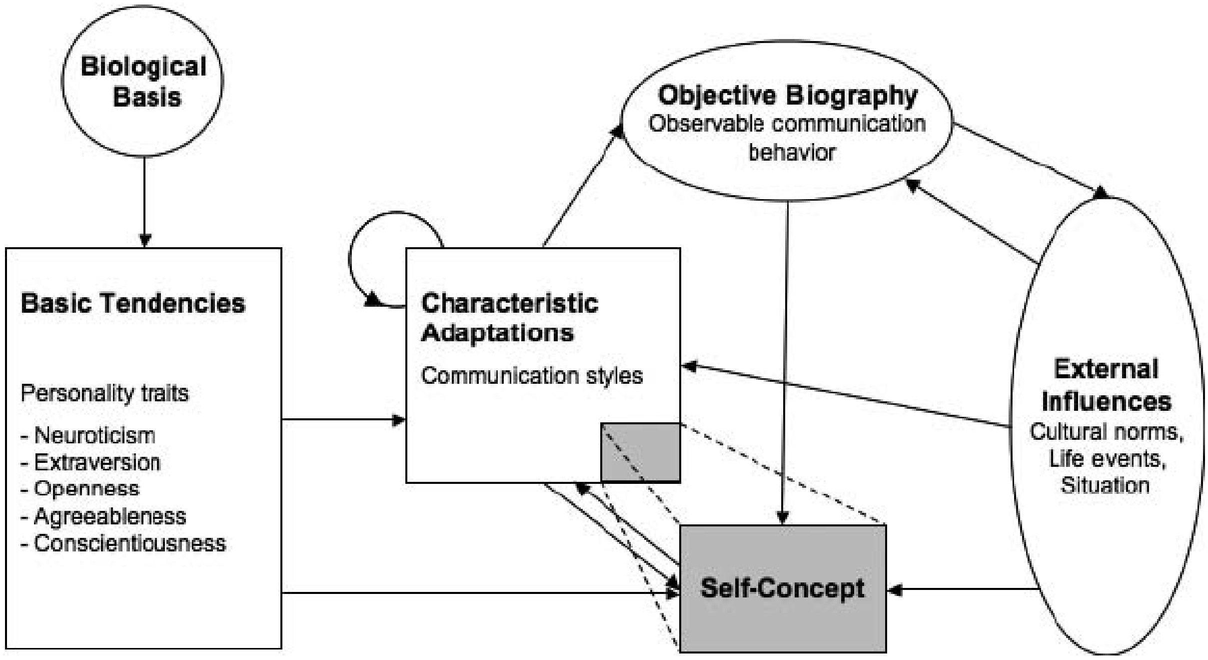
1.2. Interpretations of the Five Personality Domains

1.2.1. The Five-Factor Theory

FFM has been criticized for lack of theoretical underpinnings. Opponents have criticized some of the factors by pointing to a perceived lack of valid mechanisms (Eysenck, 1992). To answer the critics, McCrae and Costa (1999) developed the five-factor theory (FFT) to contextualize personality development in humans. The framework has generated further research targeting specific aspects of the theory. A brief presentation of the theory follows.

Figure 1

Graphic representation of the five-factor theory personality system



Note. Illustration published in Waldherr and Muck (2011), based on McCrae and Costa's (1999) framework.

In their description of the FFT, McCrae and Costa (1999) postulated that the phenotypic five domains are in fact *basic tendencies* derived from a *biological basis.* As illustrated in Figure 1, the basic tendencies interact through *dynamic processes* with each other and *external influences,* such as the social environment, to create behavior (*objective biography*). The basic tendencies are observed in *characteristic adaptations*, which are also affected by external influences. The characteristic adaptations include the *self-concept*, which is the basis of self-report. The system is interpreted both cross-sectionally and longitudinally. An example of cross-sectional interpretation is that external influence serves as a *situation,* which for example could be a guru coming to town. Basic tendencies might be tendencies to high *openness to values*, the characteristic adaptation being an interest in religion and the objective biography being Astrid joining the Hare Krishna movement.

In their ontological interpretation, McCrae and Costa (1999) stressed the distinction between basic tendencies and characteristic adaptations, where the former has been described as "abstract psychological potentials" and the latter as "their concrete manifestations" (p. 143). Etiologically, the five dimensions are considered to be basic tendencies, meaning they originate from biological bases and are not affected by environmental experiences. This means that the phenotypic expression of human variability that is measured by five-factor instruments are reflections of a biological variability in the human species. This assumption has been sought justified by studies reporting universality, stability, high degree of heritability, and low shared environmental effects in personality traits.

1.2.2. The Lexical Interpretation

Contrary to McCrae and Costa's (1999) interpretation, Saucier and Goldberg (1996) interpret the five factors as merely a taxonomy. They have proposed maintaining a strict distinction between the *genotype* (i.e., underlying causal properties) and the *phenotype* (i.e., observable characteristics), where the nature of the five factors are only known at the phenotypic level. Saucier and Goldberg (1996) have described their ontological viewpoint with a quote from John and Robins (1994) about the lexical perspective, which "(...) makes no explicit assumptions about the ontological status of traits or about the causal origins of the regularities to which they refer" (p. 138). In that regard, they have opposed the term "trait" due to its implied stability and causality, proposing the word "attribute" as a better description of the five factors. Saucier and Golberg (1996) have even stated that: "(...) [if they] were in charge of the world, [they] would ban the use of the term 'theoretical'" (p. 22), implying that a theory about the five factors' etiology, as the FFT represents, is superfluous.

We are left with two different interpretations of the Big Five dimensions' ontology and etiology. The interpretation inferred from the FFT emphasizes the biological underpinnings of the five factors and leads to hypotheses about a high degree of universality and heritability as well as a belief that the factors are creating behavior—making them *causal*. In contrast, the proponents of a *lexical interpretation* hypothesize cross-cultural variation and make no assumptions about heritability, believing the factors to be descriptions only. They make no assumptions about the etiology of the factors, while proponents of the FFT believe they are broad biological factors made up of more specific biological factors (i.e., facets)—that the phenotypic expression reflects the genetic architecture. A review of the literature regarding aspects of these hypotheses follows.

1.3. How Universal are the Big Five?

The universality of the FFM is a core argument for the biologically-dominant etiology of personality and the five basic tendencies proposed in the FFT. A cross-cultural and ubiquitous dimensionality would imply endogeneity—that the basic tendencies are derived from a universal biological basis. The issue of universality is extensively researched and debated. Fetvadjiev and van de Vijver (2015) have distinguished between three lines of investigation regarding the universality of the Big Five: i) the *etic* approach: questionnaire studies seeking to replicate FFM across cultures, ii) the *emic* approach: psycholexical studies exploring indigenous models of personality, and iii) *combined* etic-emic approaches: studies that examine personality structure with a variety of methods to assess the overlap between indigenous models and the FFM.

Translated versions of the NEO PI-R and the NEO Five-Factor Inventory (NEO-FFI) have replicated the five-factor structure in 50 and 56 cultures, respectively. Both NEO PI-R and NEO-FFI demonstrate a high degree of replicability in Western cultures (McCrae & Terracciano, 2005; Schmitt et al., 2007). Of 50 cultures assessed with NEO PI-R, 49 have been replicated with acceptable congruence measures. Openness to experience has been the least replicated domain, reporting acceptable congruence in 41 of the 50 cultures (McCrae & Terracciano, 2005). Measures of internal validity have been reported cross-culturally, implying universal personality traits by displaying similar patterns of retest reliability, cross-observer agreement, stability, and heritability of facets (McCrae et al., 2011).

Although these results point toward unequivocal universality, the notion has been nuanced by etic, emic, and combined approaches. Further investigation of etic research has revealed that factor loadings have been consistently lower in less-developed cultures (Smaldino et al., 2019). This raises hypotheses and questions about the universality of personality dimensions.

Emic approaches of other languages have yielded solutions other than five factors, supporting a lexical interpretation of the factors. Saucier and Goldberg's (2001) review of emic studies reported five-factor structures resembling the English Big Five structure in German, Dutch, Polish, Czech, and Turkish samples. However, different dimensionality was reported in Italian, Hungarian, Korean, Hebrew, and Filipino samples.

An alternative to the FFM, the HEXACO model has been developed with cultural differences in mind (Ashton et al., 2004). Whereas FFM is based on the English language, the HEXACO model is originally based on seven different languages (Dutch, French, German, Hungarian, Italian, Korean, Polish). Five of the six HEXACO factors resemble the FFM's five factors, with an additional *honesty-humility* factor. The convergence with the Big Five personality dimensions seems supportive of the interpretation proposed in FFT. However, the emergence of the factors in HEXACO might be due to a relative Western and industrialization bias. To correct for the potential bias, Saucier et al. (2014) included Chinese, Filipino, Turkish, Greek, Maasai, and Senufo in addition to Polish, Hungarian, and Korean in a lexical bottom-up approach similar to the development of HEXACO and the FFM. They reported a quite different solution—a two-factor solution of *social self-regulation* and *dynamism*. However, social self-regulation correlated with agreeableness and conscientiousness, and to some degree with neuroticism. Dynamism correlated with extraversion. Both the FFM and HEXACO openness to experience constructs hardly correlated with either of the two factors.

Evidence from a culture even farther from Western culture has been provided by Gurven et al. (2013). They used the etic approach and failed to replicate FFM in the Tsimane forager-horticulturalist men and women of Bolivia, where 11 factors were retained. A two-factor solution was proposed nonetheless, due to parsimony and internal consistency. Gurven et al. named the factors *pro-sociality* and *industriousness*. A two-factor solution has also been suggested in Western cultures. Digman (1997) proposed two factors overarching the NEOAC: *Alpha* (α) and *beta* (β). α consists of A, C, and a reversed N factor, while E and O constitute β. McCrae and Costa (2008) opposed the existence of α and β. They argued that the overarching factors were merely due to observer bias (i.e., negative and positive valence) pointing to the emergence of a two-factor solution in samples based on peer-report, while five factors are retained in samples based on self-report.

The universality of the Big Five therefore seems ambiguous. Within the etic approach, the factors are extensively replicated, except for openness to experience. The emic and etic-emic approaches to replication have, as explained above, yielded varying results. The claim of endogeneity can not be inferred solely from the splayed universality literature.

1.4. How Valid and Heritable are the FFM Facets?

In their description of the FFT, McCrae and Costa (1999) postulate that the five dimensions are broad domains made up by more narrow and specific facets. Within the FFT framework, the facets should not be understood only as partitioned variance within a factor, as hypothetically would be the case within a strictly lexical interpretation, but as building blocks for the factors. Appraising the facets therefore generates valuable information in exploring the ontology of the Big Five dimensions. Estimating the *heritability* of the facets also yields insight into the etiology of the aforementioned domains by investigating their proposed biological origin.

In order to assess the facets, a recapitulation of the facet creation debate and measures of internal consistency is necessary. The facets were constructed by Costa and McCrae (1995) for the purpose of specificity. They added six facets to each domain (see Table 1). By diverging from the original simple structure, the addition of facets has increased the predictive ability of the model (Paunonen & Ashton, 2001). However, the increased complexity has also escalated the number of cross-loadings. As an example, several extraversion facets have been shown to load on the E, A, and C factors (Costa & McCrae, 2008). This weakens the discriminative validity of the constructs. Costa and McCrae's intention was to identify mutually exclusive clusters of closely covarying elements within each domain (1995, p. 25). Yet, the conceptualization of the facet structure has been sparsely described. Boyle (2008) described the process as driven by theoretical insight and intuition. Contrary to the empirically founded domains, the facets were predominantly theoretically derived. Block (1995) characterized the method bluntly as "intelligent arbitrariness" (p. 201)—not rooted in factor analysis, formal theorizing, or ineluctable empirical findings. Admitting to arbitrariness, Costa and McCrae (1995) argued certain subdivisions of the domains to be more meaningful than others.

Empirical scrutiny of internal validity and construct homogeneity has revealed a wide range of Cronbach's alphas, which have varied as a function of sample. For example, a mean of .70 in the range of .56-.81 has been reported in the U.S. normative sample (Costa & McCrae, 1992). In a sample of French military personnel, the mean was .63, ranging from .31 to .77 (Rolland et al., 1998). Analysis of the internal consistency of the current sample was estimated by Røysamb et al. (2018), who reported a range between .47 and .85 and a mean of .67.

Another property of the facets that might shed light on the etiology and ontology of the Big Five, are the *heritability* estimates. Heritability refers to the proportion of variance in a phenotypic trait that is due to variation in genetic factors (Plomin, 2013). Heritability is estimated in family studies from the correlation between phenotypic and predicted genetic resemblance between family members (Fisher, 1919; Wright, 1921). The heritability statistic (h^2) is stated between 0 and 1, where 1 indicates that all phenotypic variance is due to genetic factors. The variance that is not accounted for by genetic factors is due to environmental variance and measurement error. The heritability of personality has been thoroughly assessed. In general, heritability of personality traits has been estimated to be about 40 % (Vukasović & Bratko, 2015). The heritability of the NEOAC domains has been reported to be in the .51-.58

range (Loehlin et al., 1998). The same has been reported for the similar HEXACO dimensions (Kandler et al., 2019).

Compared to physical properties such as height (h^2=.69-.93) (Silventoinen et al., 2003) and psychiatric conditions such as schizophrenia (h^2=.73-.79) (Hilker et al., 2018), the heritability estimates of the five factors are less prominent. This indicates substantial environmental contributions to personality development. In a meta-analysis of behavioral genetic research on the development of personality, Briley and Tucker-Drob (2014) reported increasing phenotypic stability throughout life, with relative stability achieved already in early adulthood. An important question is what contributes to stability. Briley and Tucker-Drob reported increased genetic influence on phenotypic stability until early adulthood. The effect then remained constant throughout life, explaining 75 % of the stability. Family environment contributed the most to stability in early childhood but decreased as the genetic effect on stability increased. The unique environmental influence, on the other hand, increased steadily across the lifetime from zero to a contribution of nearly 20 % in late adulthood, leading the authors to conclude that the "(...) life-span trend of increasing phenotypic stability (...) predominantly results from environmental mechanisms" (p. 1303). Mounting evidence has emphasized the role of genes and the environment in the etiology of personality.

However, research of the facet heritability in NEO PI-R has been limited; to our knowledge, only two studies have investigated this. Similar to the alphas of the facets, the reported heritability estimates also have had a wide range: both in a Canadian sample, where estimates varied from .26 to .52 (Jang et al., 1996) and a study using largely the same Canadian sample in addition to a German sample, reporting estimates between .27 and .49 (Jang et al., 1998). The scarce replication literature creates uncertainty regarding the actual heritability of the facets.

In their description of the FFT, McCrae and Costa (1999) described the Big Five domains as products of the facets. However, the empirical evidence for the facets is limited. The existing literature has displayed a serious degree of variation in facet validity. The facet heritability literature is meagre, making the etiological estimates unsettled. From an ontological perspective, there is still uncertainty regarding whether the facets are building blocks for biologically-derived domains or merely partitioned variance.

1.5. What Genetic and Environmental Factors are Etiologically Involved in Phenotypic Dimensionality?

The Big Five phenotypic dimensions could hypothetically emerge from a range of genotypic and environmental factors. According to the FFT, they are basic tendencies that emerge from a *biological basis* (McCrae & Costa, 1999). Based on this assumption, a genotypic five-factor structure might be predicted. The environment is not assumed to systematically correlate and affect the structuring of personality development. As reviewed above, different factor structures across and within cultures have been reported, challenging the notion of a genotypic five-factor structure. Empirical dissection of such notions is possible through multivariate modeling. Phenotypic variance is decomposed into genotypic and environmental variance. Covariance matrices embedded into these models give information about the systematic variation of genetic and environmental contributions to the phenotype. The estimated *genetic correlation* is a statistic that refers to the covariance between traits that are due to genetic causes and correlation between genetic influences (Robertson, 1959). The assumption of genetic correlation is derived from the mechanism *genetic pleiotropy*, which occurs when two or more traits are affected by a single gene. In the current context, this might manifest itself as traits, for example as neuroticism and extraversion partially sharing a genetic basis.

Using translated versions of the NEO PI-R, Yamagata et al. (2006) examined genetic and environmental correlations between personality facets in twins from Canada, Germany, and Japan. In line with the FFT postulates, five genetic factors were evident in all samples. Only four environmental factors emerged in Japan and Germany, while five were evident in Canada. To answer whether the genetic and environmental factors were the original five factors, congruence coefficients were estimated. N, E, O, A, and C were replicated in every sample phenotypically. Interestingly, the five factors were more clearly identified as N, E, O, A, and C in the genetic factors than in the phenotypic. The authors believed the distinctiveness of the additive genetic factors to be strong evidence for the universality and endogeneity of the Big Five. They suggested that genes structure the expression of personality.

McCrae et al. (2001) reported similar results in a preceding study. It should, however, be noted that this must be interpreted with caution due to severe methodological shortcomings. Interesting nonetheless is the application of correction for *implicit personality theory*, i.e., the tendency of humans to cluster personality traits, to the environmental factors.

After controlling for implicit personality theory, they reported *true environmental variance* to load on two factors they named "love" and "work." The authors reached the conclusion that the phenotypic structure reflects the genotypic, while the environmental factors are not sufficient to structure phenotypic personality.

The genetic structure was also examined by Franić et al. (2014). They factor analyzed the genetic and environmental influences and presented evidence for a five-factor structure, both environmentally and genetically. Visual inspection of the factors indicated congruence with the FFM domains, pointing to the genetic and environmental contribution in structuring personality. The Big Five personality dimensions then seem to be underpinned by genetic clusters resembling the five-factor structure. Two studies in particular nuance this notion.

Jang et al. (2006) reported evidence for a one-factor solution in the same Canadian, German, and Japanese sample as Yamagata et al. (2006). Due to the ability to explain variance, Jang et al. nevertheless argued for two genetic factors. The factors are consistent with the aforementioned α and β factors, except for agreeableness, which only loads on the Canadian α factor. Concluding their work, the researchers claimed that the five domains and the two factors were etiologically heterogeneous.

The same level of scrutiny has been applied to the facets. Jang et al. (2002) identified two genetic and two environmental sources of variability in each of the five domains, in contrast to the proposed six facets. The sources of variability were indeed congruent with the *aspects* proposed by DeYoung, Quilty and Peterson (2007). Whether they constitute a level of domains or facets or a level in between remains unsettled.

The question of which genetic and environmental factors are involved in a five-dimensional phenotype remains open. Anything from one to ten genetic sources and multiple environmental sources seems feasible. To grasp the etiology of the Big Five, further investigation of the genetic and environmental variation is necessary.

1.6. How to Interpret the Big Five Personality Traits?

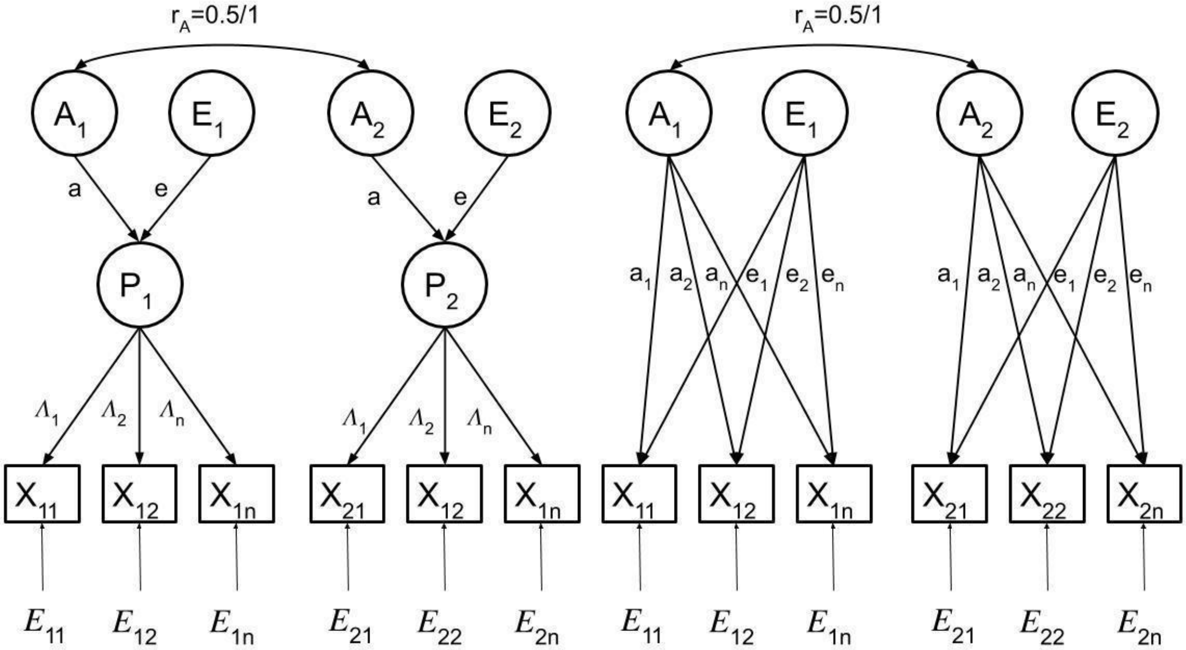
McCrae and Costa (1999) have postulated that "personality traits are endogenous basic tendencies" (p. 145) in their description of the FFT. This postulate reflects what Franić et al. (2014) refer to as a *realist interpretation*, which considers the psychological traits in question to be *veridical,* i.e., to coincide with reality, emerging from a biological basis and ultimately anchored in genes. The realist interpretation assumes that the common language reflects the psychobiological substrates in human trait variation. Neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness are not assumed to be descriptions of clusters of correlated behavior but to be entities that cause behavior (McCrae & Costa, 1999).

We call the challenging position derived from a strictly psycholexical approach *constructivist interpretation*. Proponents of a constructivist interpretation would assume that personality traits identified through factor analysis reflect semantic clusters in the language. There are, for example, many terms for being *lively* and *outgoing*, which are words that are semantically similar and might be used interchangeably. Where proponents of the realist interpretation assume this large cluster of synonyms to reflect the endogenous substrates of extraversion, proponents of a constructivist interpretation assume this cluster to reflect the human perception of phenotypic variation.

The two positions may be tested by comparing fits between so-called *independent pathways models* (IPMs) and *common pathways models* (CPMs). CPMs resemble psychometric factor models by explaining variance in a set of variables by a latent, common factor (Rijsdijk, 2005). IPMs make no such constraints on the variance, and thereby allow genetic and environmental sources of item variation to load directly on observable items. A better-fitting CPM indicates that the five factors are mediating mechanisms between genes and behavior, as veridical psychological entities are supposed to. If CPMs fit the data best, the results favor the realist interpretation, but if IPMs fit the data best, the results are unsupportive.

Figure 2

Graphic representation of independent pathways models and common pathways models



Note. Illustration of a CPM to the left and an IPM to the right. A = Additive genetic factor. E = Non-shared environmental factor. P = Psychometric factor. X = Item. *E* = Error. The abbreviations A and E are not to be confused with the abbreviations for agreeableness and extraversion, which are also “A” and “E.”

The most recent study that applied this method was performed by Franić et al. (2014). They explicitly tested the validity of the realist interpretation of the FFM by comparing how the CPM and IPM fit to the data. They reported incomplete mediation of genetic and environmental influences by the latent phenotypic factors. In other words, IPM was a better fit. The authors interpreted this as indicative of the A and E components influencing items directly, not through latent factors. Franić et al. (2014) concluded by describing the Big Five as “*statistical constructs*” rather than “*causally efficient entities*” (p. 601).

Where Franić et al. (2014) assessed the entire model, Johnson and Krueger (2004) compared CPMs and IPMs of each of the domains individually. The two information criteria that were applied yielded diverging results. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) emphasize parsimony by correcting for number of variables and sample size, but to different degrees. When BIC, which favors parsimony to a higher degree than AIC, was applied, the CPMs showed to best fit neuroticism and extraversion—indicating unity of the two domains. However, when AIC was applied, the CPMs never fit the data better than the IPMs, leaving Johnson and Krueger (2004) inconclusive about the ontology of the personality dimensions. Jang et al. (2002) have also compared CPMs to several IPMs to each domain. In their study, CPMs displayed to fit each domain worst, leading Jang et al. (2002) to conclude: "The present results suggest that higher-order traits such as "neuroticism" do not exist as veridical psychological entities (...)" (p. 99).

Similarly, several CPM and IPM comparisons have been applied to each of the HEXACO domains (Lewis & Bates, 2014). Assessing fit with the AIC, four of the six domains displayed better-fitting CPMs. However, three of the best-fit CPMs mediate only genetic contributions, leaving environmental contributions to influence the facets directly. An IPM fit conscientiousness and openness best. Lewis and Bates' results imply genetic unity in four of the six HEXACO domains, where three of them resemble the FFM domains (neuroticism, extraversion, and agreeableness).

The conflicting results from CPM and IPM comparisons remain a source of debate and call for replications. Some of the authors have drawn conclusions in favor of a constructivist interpretation while others remain hesitant regarding the question of ontology. When comparing the evidence, there are several methodological discrepancies that need to be taken into consideration, leaving the unsettled ontology indiscernible. Do the five domains comprise a veridical psychological model?

1.7. The Etiology and Ontology of the Big Five Personality Traits

The FFM is the most widely-applied model of personality in research and culture today. How the five personality traits should be understood and interpreted nevertheless remains unsettled, which has given us renewed interest in the questions regarding their etiology and ontology.

Proponents of the five-factor theory assume the five factors to be veridical psychological entities that cause human behavior. This interpretation rests upon empirical evidence showing cross-cultural replications and a high degree of heritability and unity in the dimensions. Critique of this interpretation is made with reference to culturally-sensitive studies that fail to replicate the Big Five, making the universality ambiguous. To our knowledge, no bottom-up replication of the five-factor dimensionality in a Norwegian sample has been published. In the translation and adaptation process, only results from confirmatory, not exploratory, factor analysis was published (Martinsen et al., 2011). Whether the five-factor structure emerges in an etic investigation of this sample is uncertain.

The ambiguous universality can be a reflection of only partial endogenic genetic contribution or of a potentially deficient instrument—indications of both are evident in the facet literature. However, the understanding of the supposed building blocks of the Big Five dimensions is limited. Investigation of the association between new heritability estimates and measures of construct homogeneity elaborates the discussion about the domains' etiology and ontology.

How are the supposed building blocks involved in the composition of genetic and environmental factors? According to the FFT, there should be five genetic factors, and the factors should be identical to the originally proposed phenotypic factors. Yet, authors report anything from one to ten genetic sources underpinning the Big Five personality traits (Jang et al., 2002, 2006). Contrary to what may be inferred from the FFT, empirical data indicate systematic variation in the environment that contributes to the etiology of personality. Theoretical explanations for these empirical findings are lacking in the FFT; therefore, continuing the empirical and theoretical exploration is crucial for the understanding of the Big Five personality dimensions.

However, by answering questions about the composition of the factors, we do not answer *how* the five factors work in relation to aggregations of cognition, emotion, and behavior. Whether they are merely descriptions of phenotypic personality expression or actual psychological entities that mediate the interplay between environmental and genetic contributions is still the subject of debate.

To clarify, the objective of this thesis is to investigate these ontological and etiological questions in three ways. i) We aim to appraise the facets through two steps, one step being a principal component analysis of the Big Five facets to assess the replicability of the model and to structure the phenotypic level of the etiological pathways models. The other step is an assessment of the validity of the building blocks (i.e., facets) through estimation of heritability and internal consistency and the association between them. ii) We aim to explore which genetic and environmental factors might contribute to which personality traits by applying principal component analysis to additive genetic and non-shared environmental facet correlation matrices. iii) We aim to illuminate the fundamental ontological interpretation of the Big Five by comparing three etiological pathways models: a common pathways model and an independent pathways model, both based on the phenotypic rotation, and a local etiological independent pathways model based on the structure of the non-shared environmental and additive genetic influences. The comparison will shed light on the crucial question of whether the five factors mediate environmental and genetic contributions or not, and should be interpreted as veridical entities.

2. Method

2.1. Behavioral Genetics

When investigating the nature of the FFM, this thesis and the relevant literature applies a behavioral genetic framework. A brief review of the framework is therefore deemed appropriate.

Tracing back to Mendel's laws of heredity, behavioral genetics provides a theoretical and empirical framework for estimating the contribution of genes and environment to observable traits in living creatures. Mendel's original laws lay the groundwork for the basic unit of heredity: genes (Bateson & Mendel, 1913/2013). Most psychological traits are *polygenic* (i.e., affected by many genes), which makes many additive combinations possible. This results in a normal distribution of phenotypes, such as personality dimensions. The basic principle of quantitative genetics is that genetic relatedness is correlated with phenotypic resemblance, i.e., one expects higher phenotypic resemblance in siblings than in cousins (Fisher, 1919; Wright, 1921).

A basic assumption in behavioral genetics is that the *phenotype* (P) consist of genetic contributions (Ge) and environmental contributions (En). Hence,

$$P = Ge + En$$

Twin studies identify two possible ways in which genes contribute to the phenotype: additive genetic effects (A) and non-additive effects (D). Hence,

$$Ge = A + D$$

The D component consists of epistatic effects (a phenomenon in which the effect of an allele is dependent on the presence or absence of alleles at other loci) and dominance effects (interaction effects between alleles at the same loci). According to quantitative genetic theory, A correlates 1.0 across monozygotic (MZ) twins and 0.5 across dizygotic (DZ) twins. Dominance effects, which are the main component of D, correlate 1.0 across MZ twins and 0.25 across DZ twins. However, most studies do not calculate the D parameter, because the classical twin design (CTD) does not allow for the estimation of more than three parameters. *Narrow sense heritability* (h^2) refers to additive genetic effects (a^2), and *broad sense heritability* (H^2) refers to both additive and non-additive genetic effects ($a^2 + d^2$) (Plomin, 2013).

Environmental contributions consist of shared environmental effects (C) and non-shared environmental effects (E). Hence,

$$En = C + E$$

E contains both measurement error and hypothetically true environmental variance. Any resemblance between twins can stem from being exposed to and operating in the same family (i.e., shared) environment. If rDZ > 0.5rMZ, resemblance is greater than genetically predicted and assumed to be due to C.

By assuming zero covariance between the terms and including different kinds of genetic and environmental contributions, P = GC + EC expands to:

$$\text{Var}(P) = \text{Var}(A) + \text{Var}(D) + \text{Var}(C) + \text{Var}(E)$$

2.2. Sample

The twins in this study were recruited from the Norwegian Twin Registry (NTR), established by the Norwegian Institute of Public Health (NIPH), Oslo University Hospital and the University of Oslo (Harris et al., 2006). The current sample was randomly drawn from NTR as part of a project on health complaints, personality, and quality of life (see for example Vassend et al. (2017)). The sample consists of twins born between 1945 and 1960. At the time of the data collection (2010–2011), the mean age was 57.4 years (SD=4.6), making it a sample with assumed high phenotypic stability (Briley & Tucker-Drob, 2014). The questionnaires were sent to a total of 2,136 twins. The response rate was 71 % (1516 responders) after reminders. 1272 individuals were pair responders. 244 were single responders, (i.e., 244 are missing). The question of zygosity was already determined using questionnaire items. This method has shown to correctly classify 97.6 % of the twins (Magnus et al., 1983). The cohort contains only same-sex twins: 456 MZ female twins, 523 DZ female twins, 290 MZ male twins, 247 DZ male twins. In treating missingness, we allowed half of the items in a facet to be missing. Facets containing more than four missing items were deemed *Not Available*.

2.3. Measures

The *Revised Neuroticism, Extraversion, Openness—Personality Inventory* (NEO-PI-R) consists of 240 items that form five factors with six facets each. The items consist of statements with a five-point *Likert scale*, ranging from *strongly disagree* to *strongly agree* (Costa & McCrae, 1992). The present study used the Norwegian translation, which has shown to have a facet congruence with the original American sample ranging from .85 to 1.0 (*M*=.9753, *SD*=.0297) and factor congruence ranging from .97 to .99 (*M*=.9780, *SD*=.0084) (Martinsen et al., 2011). In our sample, Cronbach's alphas for the five factors were 0.92 (neuroticism), 0.87 (extraversion), 0.87 (openness), 0.83 (agreeableness), and 0.86 (conscientiousness). Alphas for the facets ranged from 0.46 (A6 tender-mindedness) to 0.84 (N1 anxiety), with a mean of 0.67 (*SD*=0.085).

2.4. Analyses

2.4.1. Structural Equation Modeling

Structural equation modeling (SEM) is a statistical procedure widely used in the social sciences. For the most part, the field of psychology studies abstract concepts that cannot be directly measured. SEM allows for the estimation of these abstract *latent* concepts based on measured *manifest* variables. One may infer the value of a latent variable by estimating covariation in variables that are supposedly linked to the latent construct. The approach assumes that covariation between the manifest variables stems from the latent variable and not reflect a causal relationship between them (Bollen, 1989).

The SEM consists of dependent and independent variables, and the *structural model* describes the relationship between them. The variables are estimated based on the manifest variables, which belong to the *measured model*. The parameters of a SEM model consist of variances, covariances, and correlations between the latent and manifest variables, which are estimated from raw data (Bollen, 1989). Formulation of the parameters is done in terms of linear algebra, matrix algebra, and path diagrams, as demonstrated in the following section.

2.4.2. Model

Behavioral genetics assumes four sources of variance in twin traits:

$$Var(P) = Var(A) + Var(D) + Var(C) + Var(E)$$

An AE-model was opted for in our analysis, constraining the variance of C and D to zero. The decision is based on the following arguments: i) Omission of the C and D factors is close to standard in most behavioral genetic personality research due to the generally low ability to explain variance (Bouchard & Loehlin, 2001; Briley & Tucker-Drob, 2014). Studies similar to ours have opted for an AE-model (Franić et al., 2014; Jang et al., 2002, 2006; Johnson & Krueger, 2004; McCrae et al., 2001; Yamagata et al., 2006). ii) Upon inspection, the C and D components seem negligible for the majority of the facets in this thesis (see Appendix A). iii) The mean age of the sample is 57.4 years, and C components are expected to decrease with age (Briley & Tucker-Drob, 2014; Loehlin & Martin, 2001). iv) Constricting C and D reduced the number of parameters in our analysis drastically, making the model computationally feasible.

Hence, the following equation translates to the covariance matrices in Table 2.

$$Var(P) = Var(A) + Var(E)$$

Table 2

MZ variance and covariance matrix

	Twin 1	**Twin 2**
Twin 1	Var(A) + Var(E)	
Twin 2	Var(A)	Var(A) + Var(E)

Table 3

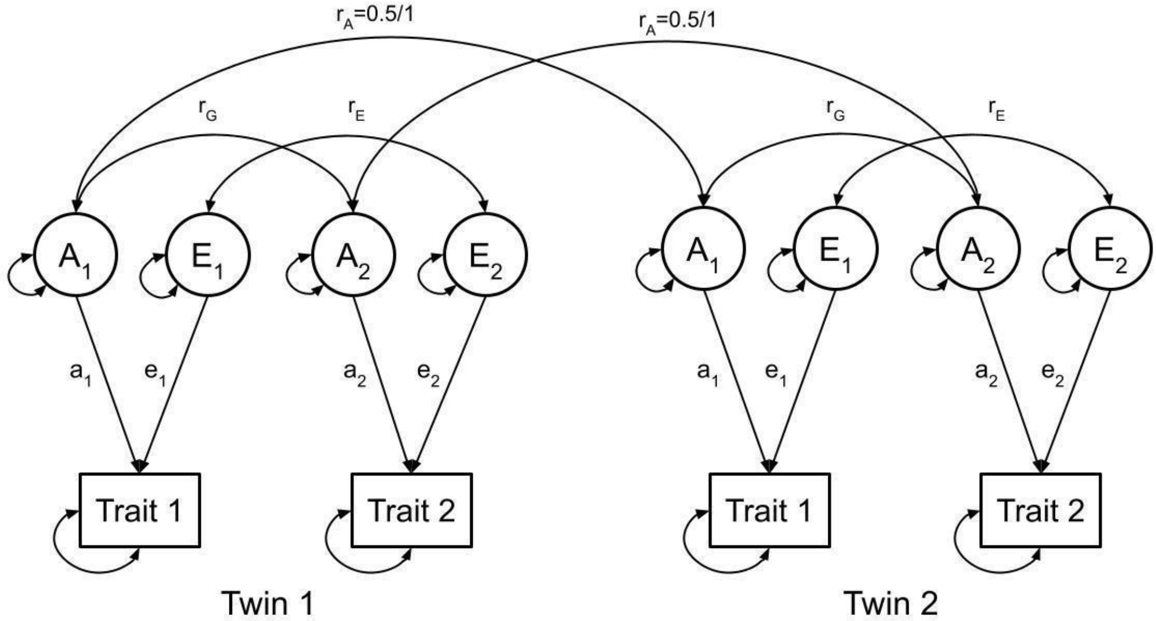
DZ variance and covariance matrix

	Twin 1	**Twin 2**
Twin 1	Var(A) + Var(E)	
Twin 2	0.5×Var(A)	Var(A) + Var(E)

A path diagram illustrates relations between each variable included in the model. The trait model describes paths between each of the observed variables (i.e., traits) and the latent additive genetic and environmental variables (Figure 3).

Figure 3

Path diagram illustrating a bivariate twin model



Formulas can be extracted from the path diagram by tracing relevant paths while following certain rules (see Wright (1934)). An example is covariation in Trait 1 between Twin 1 og Twin 2. Following the paths, the covariation consists of: $a_{1\times}1/0.5\times a_1$. The paths can be described as linear algebra organized in a matrix, as in Table 4.

Table 4

Variance-covariance matrix for a bivariate twin model

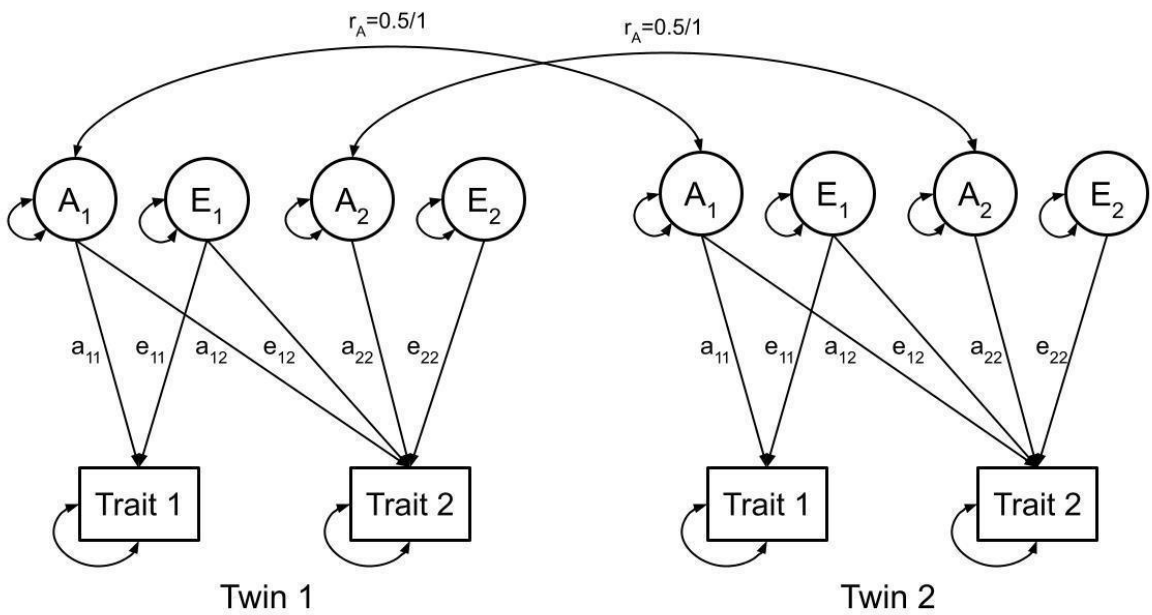
	Twin 1 Trait 1	**Twin 1 Trait 2**	**Twin 2 Trait 1**	**Twin 2 Trait 2**
Twin 1 Trait 1	$a_1^2 + e_1^2$			
Twin 1 Trait 2	$a_{1\times}r_{G\times}a_2 +$ $e_{1\times}r_{E\times}e_2$	$a_1^2 + e_1^2$		
Twin 2 Trait 1	$a_{1\times}$ $1/0.5 \times a_1$	$a_2 \times 1/0.5 \times r_G \times a_1$	$a_2^2 + e_2^2$	
Twin 2 Trait 2	$a_{2\times}1/0.5\times r_{G\times}a_1$	$a_{1\times}$ $1/0.5 \times a_1$	$a_{1\times}r_{G\times}a_2 +$ $e_{1\times}r_{E\times}e_2$	$a_2^2 + e_2^2$

A simple bivariate model as described above has quite a few parameters (paths) describing the relationship between two traits. A similar multivariate model describing the relationships between the proposed model consisting of 30 facets increases the number of parameters exponentially—to a severely computationally-demanding amount. Cholesky decomposition is therefore applied to reduce the number of parameters considerably without notably compromising the model's fit.

More specifically, the Cholesky decomposition divides a symmetrical matrix (*Z*) into a lower triangle (*T*) and its transposed upper triangle (*T'*), which is algorithmically expressed $Z=T\times T'$. The application relies on the assumption of symmetry in the relationships between the latent variables and the measured variables. In other words, the effect of A_1 to Trait 2 equals A_2 to Trait 1. Only one of the parameter triangles is used in the computation. The other triangle of parameters is removed while keeping the diagonal. This reduces the number of parameters by approximately 45 %. The number of parameters in our proposed model of 30 latent A and E variables and 30 traits is thereby reduced by 840.

Figure 4

Path diagram illustrating Cholesky decomposition of a bivariate twin model



Certain components of the model were of interest to the analysis. Phenotypic, additive genetic (A matrix), and environmental (E matrix) correlation matrices were therefore extracted.

2.4.3. Model Fitting

To execute the following analysis, the OpenMx package in R was applied (for further description, see Neale et al., 2016).

Maximum likelihood estimations were used to estimate the structural model's fit to the measured observations. In the multidimensional plane of parameters, there are several statistically likely solutions. Maximum likelihood estimations iteratively seek to maximize the likelihood that the observed data are true, given the structural model. Due to the existence of infinite iterative possibilities, a limit must be set. This creates a probability of incidentally settling on a suboptimal solution. To address this risk, several iterations were performed. This function uses the parameter estimation from the previous attempt as starting values, and repeats the process until the parameter estimations equal the previous attempt.

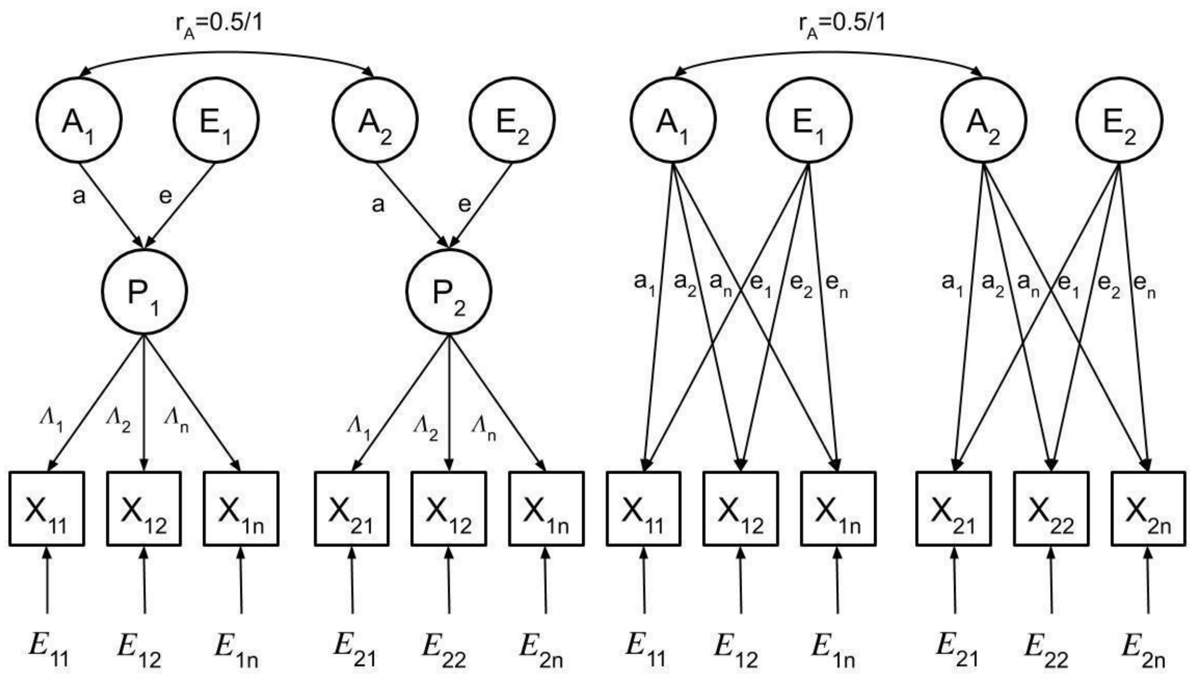
The model-fitting procedure has two goals, which represents a dilemma. Next to explaining as much of the observed data as possible, it seeks to minimize the number of parameters to increase parsimony. Information criteria such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) emphasize parsimony by correcting for number of variables and sample size. Because simpler models are less vulnerable to error, models assessed by information criteria have a higher degree of generalizability. In this study, AIC was used to compare and assess which model fit the data best. Advantages and disadvantages of the AIC are reviewed in the discussion.

Two models were proposed. The least constricted model assumed similar heritability in both sexes, while allowing differences in variation (i.e., dimension means) between the sexes. This model has 1020 parameters. The second model assumed no difference between the sexes and therefore had the fewest parameters (990). Sex differences in facet scores might also affect the validity of the proposed models. To investigate the significance of sex differences, two models were compared: i) a *sex difference* model that permits mean facet values to differ across sexes and ii) a *no sex difference* model that constrains variance across sexes. The sex difference model fit the data significantly better (*AIC diff.* = 110, $p < 0.01$). This reflects a difference in facet means across sexes. The sex difference model was therefore retained and applied in our analysis.

2.4.4. Common Versus Independent Pathways Models

Figure 5

Graphic representation of independent pathways models and common pathways models



Note. CPM to the left and IPM to the right. The a and e pathways are constrained by the psychometric factors in the CPM, while they can vary independently of one another in the IPM.

To investigate the mediating effect of five-factor constructs, a CPM was compared to an IPM. The CPM models a latent variable mediating the contribution of A and E. The IPM does not impose similar constraints. Direct influence from A and E factors on the observed variables are modelled. The CPM and the IPM are compared by applying AIC as a measure of fit. The applied CPM and IPM were based on the phenotypic structure estimated by principal component analysis. The local etiological IPM was based on the structures estimated in the A and E matrices, resembling a type of model proposed in Franić et al. (2013), and its fit was also assessed.

2.4.5. Principal Component Analysis

Principal component analysis (PCA) was applied to the phenotypic, additive genetic, and nonshared environmental correlation matrices. PCA assumes all observed variation to be accounted for by all components. Components that reflect the most variation of the variables are retained as factors (i.e., facets).

Parallel analysis is applied to determine the number of retained factors. The criteria is based on randomly-generated Eigenvalue corrected for the number of variables and sample size. Factors with Eigenvalue above the criteria are retained. Mounting evidence from simulation and comparative studies suggests that parallel analysis is more accurate in the retention of factors than e.g., the classic *Eigenvalue > 1 criteria* (see Hayton et al. (2004)).

Oblique rotation was employed to rotate the factors in all three matrices. Oblique rotation assumes correlation between the factors. We assumed correlation for factors extracted from all three correlation matrices: i) Oblique rotations are overall better fitting than orthogonal rotations when estimated in the phenotype with confirmatory factor analysis (McCrae et al., 1996); ii) *Genome-wide association* studies indicate that there are few candidate genes for specific domains (e.g., van den Berg et al. (2016) for extraversion; Smith et al. (2016) and Okbay et al. (2016) for neuroticism). This can be interpreted to reflect that a majority of genes influence multiple domains, which would be reflected in correlations between genetic factors; iii) As opposed to the phenotypic and additive genetic correlation matrices, there were no conspicuous reasons to assume the absence or presence of covarying environmental factors. We deemed it logical and parsimonious to also apply oblique rotation to the nonshared environmental correlation matrix. Similar studies also apply oblique rotation to the E matrix (Franić et al., 2014; McCrae et al., 2001; Yamagata et al., 2006).

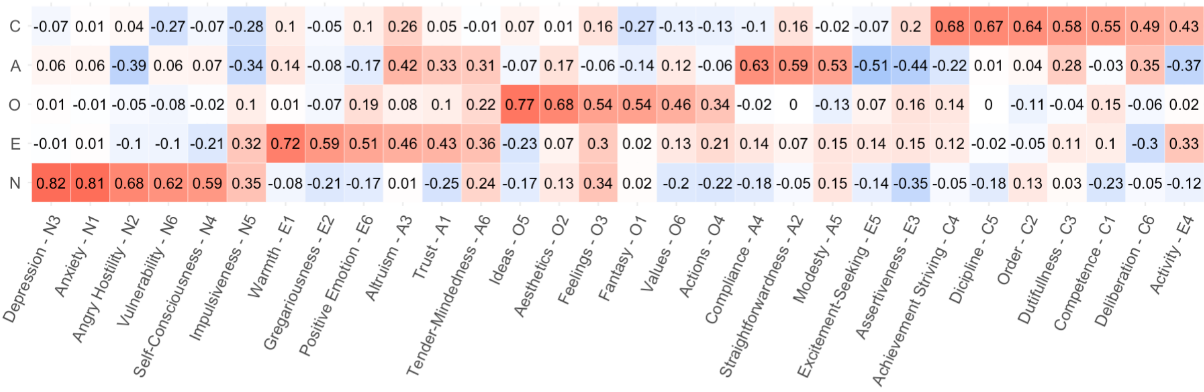
3. Results

3.1. Component Analysis of Phenotypic Correlations

Figure 6 shows that the N, O, and C factors resembled the original five-factor structure. Contrary to the original structure, the A and E domains were not distinct. The first A/E (labeled as E in Figure 6) factor had its strongest loadings from E1 warmth, E2 gregariousness, and E6 positive emotion. A3 altruism, A1 trust, and A6 tender-mindedness had their highest loadings on the first A/E factor, while they also loading relatively high on the second A/E factor (labeled as A in Figure 6). The second A/E factor had its highest loadings from A4 compliance, A2 straightforwardness, and A5 modesty, and negatively from E5 excitement-seeking. E3 assertiveness had its highest (negative) loading on this factor, but also had a moderate negative loading on N.

Figure 6

Oblique rotation of the phenotypic correlation matrix



Note. Five factors were retained when the parallel analytic criteria were applied. The scree plot is presented in Figure 14 in Appendix B.

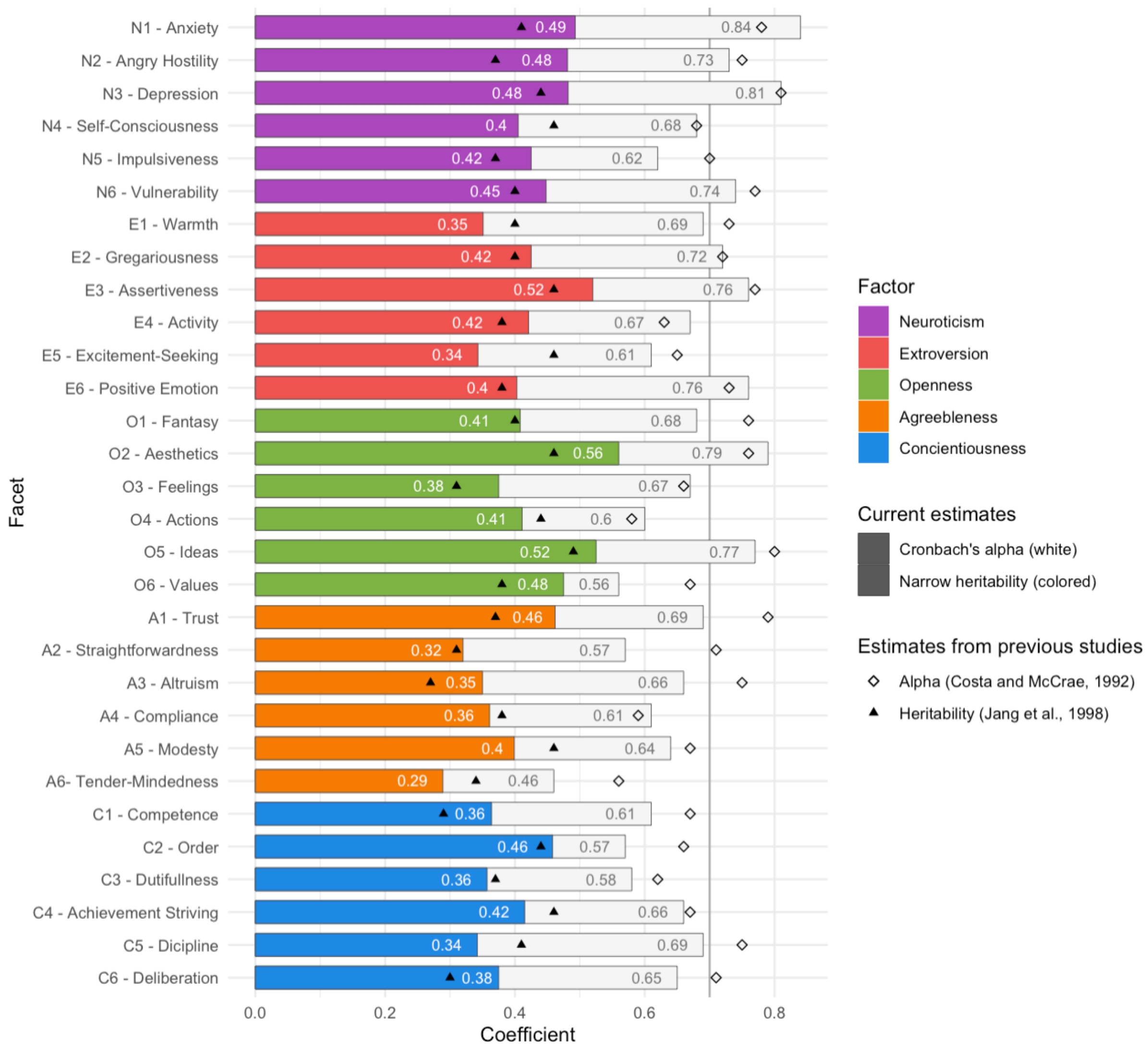
Osborne (2014) reported communality of .30 and .40 to be the most common saliency criteria in studies that identified an a priori criterion for cut-off in PCA solutions. None of the facets in the rotated phenotypic matrix loaded < .30 on any factor, but three (A6 tender-mindedness, O4 actions, N5 impulsiveness) loaded < .40 on a factor. Following Osborne's notion of "logical examination of the pattern in a factor" (p. 41), some facets seemed less "pure." Two facets (N2 angry hostility, A3 altruism) loaded > .40 on two factors. Seven facets (E4 activity, C6 deliberation, N5 impulsiveness, E3 assertiveness, O3 feelings, A1 trust, A3 altruism) loaded > .30 on two factors. Of the seven facets, three (E4 activity, N5 impulsiveness, O3 feelings) loaded > .30 on three factors.

3.2. Heritability and Internal Consistency

Heritability and Cronbach's alpha estimates for the facets displayed a wide range and correlated moderately (Pearson's *r*=.67). Figure 7 shows heritability estimates and Cronbach's alphas of each facet, contextualized with similar measures from previous studies. Costa and McCrae (1992) reported a median Cronbach's alpha of .75, while our sample yielded a median of .67. The domain alphas were considerably more robust (*M*=.87, *SD*=.03).

Figure 7

Illustration of narrow heritability and Cronbach's alpha estimates for facets in the current sample



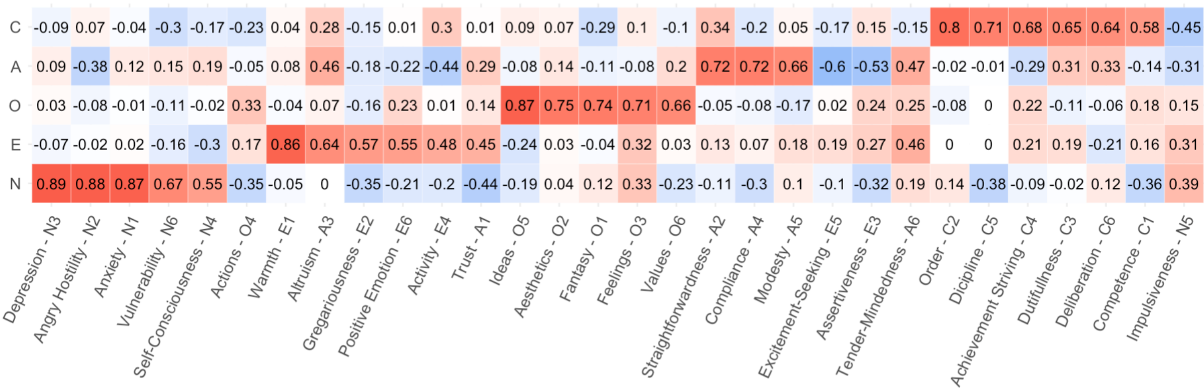
3.3. Component Analyses of Genetic and Environmental Correlations

3.3.1. Additive Genetic Correlations

Figure 8 shows that the N, O, and C constructs resembled the original and the phenotypic structures. Some unproposed loadings were nonetheless present: O4 actions loaded highest on the N factor, and N5 impulsiveness loaded highest on the C factor. Note that both facets had > .3 on two or more factors, where the original factor was of them. Agreeableness and extraversion did not bear resemblance to the original structure. However, the resemblance was closer than in the phenotypic solution. Of six proposed E facets, four had their highest loadings on E. Two of the A facets also loaded highest on the E factor. The opposite was true for the A factor, which had its highest loadings from four A facets and two E facets.

Figure 8

Oblique rotation of the additive genetic correlation matrix



Note. Five factors were retained when the parallel analytic criteria were applied. The scree plot is presented in Figure 15 in Appendix B.

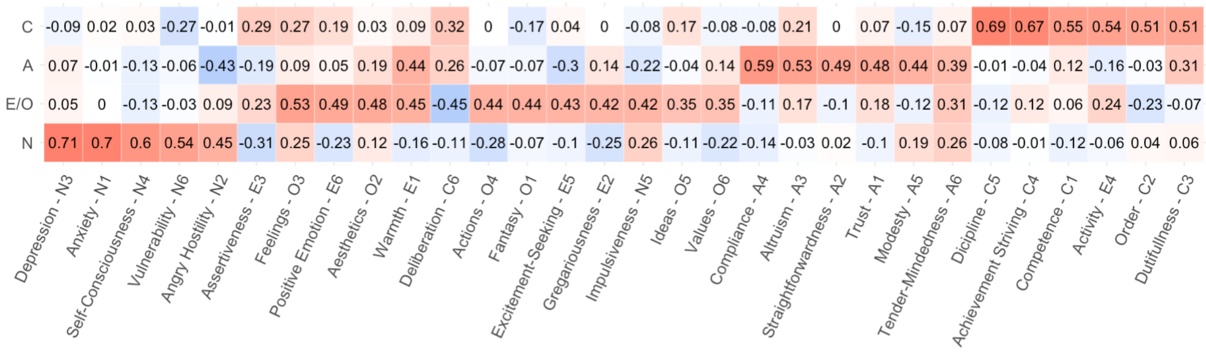
None of the facets in the rotated A matrix loaded < .30 on any factor, but O4 actions loaded < .40 on any factor. Four facets (A3 altruism, A1 trust, E4 activity, A6 tender-mindedness) loaded > .40 on two factors. Eighteen of the facets (N2 angry hostility, N6 vulnerability, N4 self-consciousness, O4 actions, A3 altruism, E2 gregariousness, E4 activity, A1 trust, O3 feelings, A2 straightforwardness, A4 compliance, E3 assertiveness, A6 tender-mindedness, C5 discipline, C3 dutifulness, C6 deliberation, C1 competence, N5 impulsiveness) loaded > .3 on two facets. Of the eighteen facets, O3 feelings loaded > .3 on three and N5 impulsiveness on four.

3.3.2. Non-shared Environmental Correlations

Figure 9 shows that the N, A, and C factors resembled the original five-factor structure. All N facets, apart from N5 impulsiveness, as well as E3 assertiveness (negatively), loaded highest on the N factor. All A facets loaded highest on the A factor. Apart from C6 deliberation, all C facets and E4 activity loaded highest on a common factor. Undoubtedly, the most interesting factor is the mixed E/O factor, which also contained N5 impulsiveness (positive) and C6 deliberation (negative).

Figure 9

Oblique rotation of the nonshared environmental correlation matrix



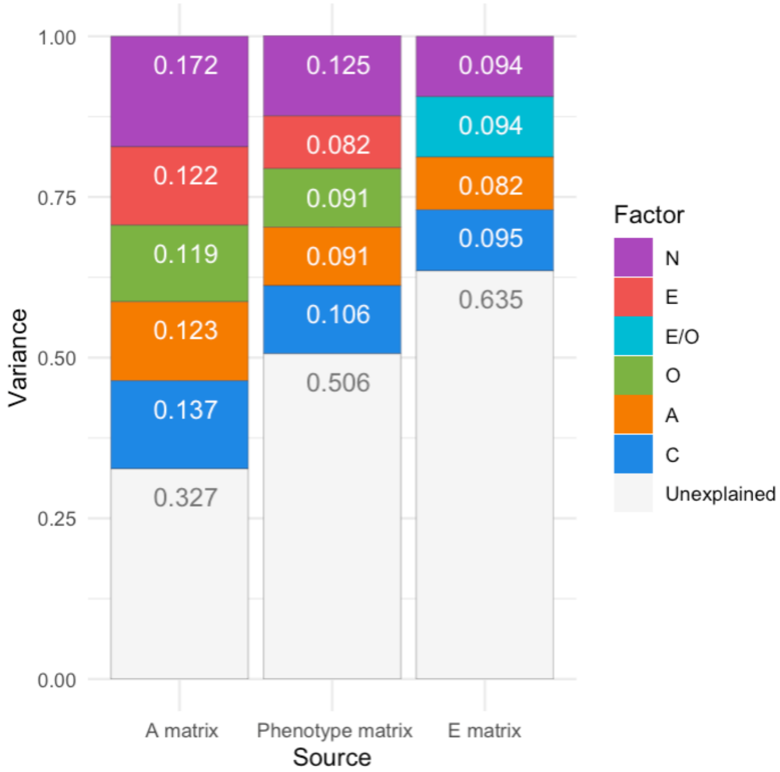
Note. Four factors were retained when the parallel analytic criteria were applied. The scree plot is presented in Figure 16 in Appendix B.

None of the facets in the rotated E matrix loaded < .30 on any factor. Four facets (E3 assertiveness, O5 ideas, O6 values, A6 tender-mindedness) loaded < .40 on any factor. Two facets (N2 angry hostility, E1 warmth) loaded > .40 on two factors. Six of the facets (N2 angry hostility, E1 warmth, C6 deliberation, E5 excitement-seeking, A6 tender-mindedness, C3 dutifulness) loaded > .3 on two facets. Of these six facets, none loaded > .30 on more than two.

The proportion of variance that was explained by the factors varied across different levels of the analysis. Figure 10 shows that the additive genetic correlation matrix explained 67.3 % of the variance in self-reported behavior, clearly more than the phenotypic matrix (49.6 %) and the non-shared environmental matrix (36.5 %). Note that the phenotypic and the non-shared environmental matrices contained systematic and unsystematic error as well. Except for neuroticism, which explained considerably more variance in the A matrix and the phenotypic matrix, the factors were quite similar in their explanatory abilities.

Figure 10

Proportion of variance explained by the retained factors



Note. A matrix = Variance from additive genetic correlation matrix. E matrix = Variance from nonshared environmental correlation matrix.

3.4. Common and Independent Pathways Models

Table 5 shows that the AIC of the IPM was significantly lower than that of the CPM. The better-fitting IPM points toward direct influences from the A and E factors, as opposed to A and E influences mediated by latent factors. The influence of the five additive genetic factors and four nonshared environmental factors in the local etiological IPM are illustrated in Figure 11. The comparisons of the two IPMs provided a lower AIC for the local etiological IPM.

Table 5

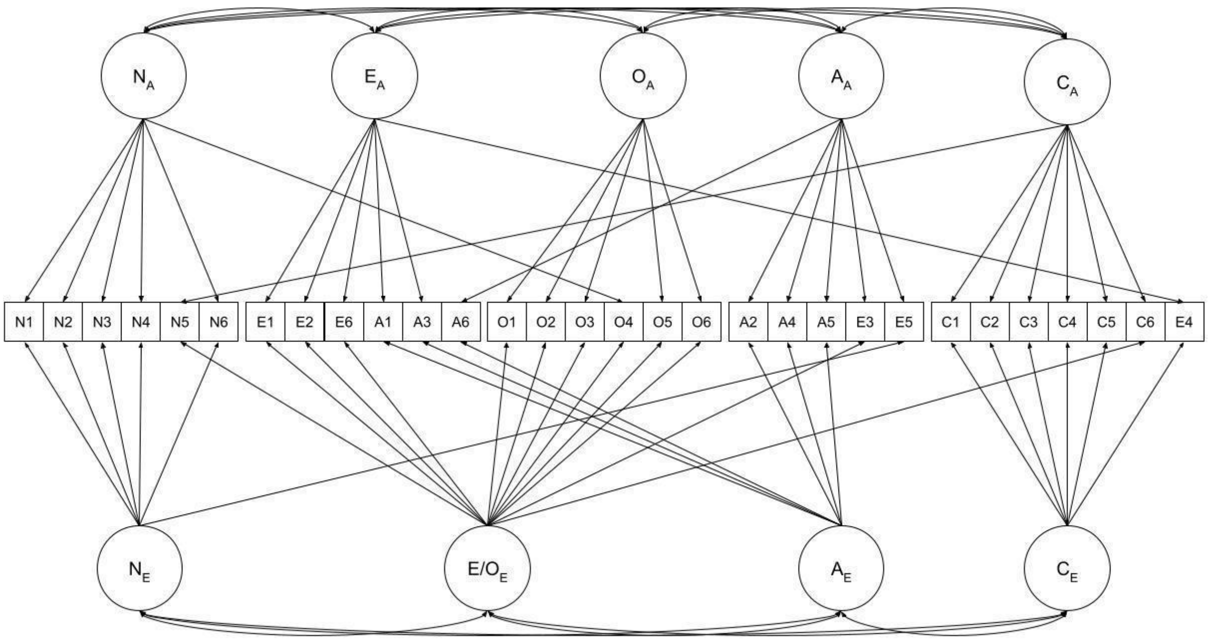
Results from the CPM and IPM comparisons

Model	Parameters	AIC
IPM	198	-43903
CPM	169	-43583
Local etiological IPM	186	-44039

Note. The CPM and the IPM were based on PCA of the phenotypic correlation matrix. The local etiological IPM was based on PCAs of the A and E correlation matrices. The difference between IPM and CPM was significant (p<0.01).

Figure 11

Illustration of the local etiological IPM.



Note. The arrows are based on PCAs of A and E correlation matrices. Highest loadings are displayed and included in the model. The facets are arranged according to the PCA of the phenotypic correlation matrix, alphabetically and numerically ordered within each factor.

4. Discussion

In the Norwegian sample, five factors emerged from the principal component analysis of the facets. The N, O, and C factors resembled the original Big Five structure to a great degree, while the A and E factors diverged considerably. The facets displayed a wide range of heritability and Cronbach's alpha estimates, which were moderately correlated. The facets also displayed large amounts of cross-loadings in every level of the analysis. Five factors were extracted from the PCA of the additive genetic contributions to the phenotypic structure. They resembled the phenotypic factors, but were also more aligned with the original FFM structure. Only four factors emerged from the PCA of the non-shared environmental contributions. The N, A, and C factors were recognizable, while E and O were combined to one factor. The local etiological IPM (Figure 11) that was constructed on the basis of the five genetic and four environmental factors fit the data better than an IPM and a CPM that were solely based on the phenotypic factor solution.

The following discussion further elaborates and contextualizes these findings. We begin with methodological discussions of the sample, the classical twin design, and the choice to apply PCA instead of exploratory factor analysis. The discussion then turns to the phenotypic factor structure, focusing on the issue of cross-cultural replicability, especially of agreeableness and extraversion, in light of translation, cultural, and linguistic issues. The validity of the Big Five facets are then discussed from a psychometric and etiological viewpoint.

The second part of the discussion revolves around the etiological influences and the fundamental ontological interpretation of the Big Five traits. We discuss how to interpret the endogenous and exogenous contributions to phenotypic personality, and which implications the diverging results might have for psychometry, genetic research, and the understanding of personality development. Finally, we discuss and contextualize the findings of the comparison between the CPM and the IPMs. We point to statistical and methodological reasons for the superiority of the IPMs and discuss how personality traits derived from the lexical approach might be interpreted.

4.1. Methodological Considerations

4.1.1. Sample

Our first consideration is the age group of the sample. Given the current sample's mean age of 57.4 years (*SD*=4.6), a high phenotypic stability can be expected (Briley & Tucker-Drob, 2014). The results may therefore contain less unsystematic error than would samples with younger twins. At the same time, the narrow age range might represent cohort effects that impede the generalizability of the results. The excluded C effects have a stronger influence in younger samples, while contributions from the non-shared environment on personality typically increase with age. One could therefore expect structures in E effects to resemble more closely the phenotypic factor structures in this sample than in samples with younger participants.

4.1.2. Possible Shortcomings of the Classical Twin Design

Before discussing the multivariate genetic results, certain considerations of the classical twin design (CTD) need to be addressed. The CTD allows only three of the A, C, D, and E effects to be estimated simultaneously (Keller & Coventry, 2005). This "inherent indeterminacy" leads to consistent biases in CTD estimations (Keller & Coventry, 2005, p. 201). The exclusion of both nonadditive genetic (D) and common environmental (C) estimations from our model left our results even more prone to undifferentiated biases. In addition to the excluded C and D effects, Keller and Coventry reported that the remaining A and E estimates may be confounded by i) gene-by-environment correlations and interactions, ii) violations of equal environment assumption, iii) violations of the assortative mating assumption, and iv) sibling interaction effects.

Gene-environment correlations are correlations in sources of trait variation and are defined by three separate processes. i) Passive correlation: Children share genes with their parents while also being exposed to the environment provided by these parents, ii) active correlation: Individuals choose environments based on genotypic inclinations, and iii) evocative correlation: Genotypic tendencies elicit responses from the environment. Gene-environment interaction refers to interaction effects between specific genetic and environmental factors. A classic example by Caspi et al. (2002) illustrates how a specific polymorphism moderates effects of early-life maltreatment on antisocial behavior. Yet, such effects are shown to be rare and seldom replicated (Munafò et al., 2009). The coarseness of the CTD design does not yield information about gene-environment correlations or gene-environment interactions. Therefore, awareness of the possible correlations and interactions must be incorporated into the interpretation.

The equal environment assumption (EEA) is another debatable presumption. In behavioral genetics, it is assumed that environmental correlation across MZ twins and DZ twins is equal, i.e., upbringing of DZ twins is equal to upbringing of MZ twins. A significant concern regarding the EEA is the greater potential for variability in the environment of DZ twins. For example, DZ twins might be of different sexes, look quite different (Tellegen et al., 1988), and share fewer childhood memories (Borkenau et al., 2002). However, there are studies that have examined differences between MZ and DZ twins that support EEA. For example Kendler et al. (1993) reported that twins who had been misinformed about their zygosity had the same degree of similarity as correctly informed twins (i.e., perceived zygosity did not influence resemblance in twins). The discrepancy between the potential for difference and measured difference was explained by Kendler et al. as MZ twins *creating* a more similar environment than DZ twins. The explanation has been criticized for tautology on the grounds of assuming heritable traits to begin with (Joseph, 1998). Although there is general consensus of the validity of the EEA, our study recognizes that possible violations of EEA might inflate heritability estimates (Felson, 2014; Plomin, 2013), which in turn would bias the conclusion in favor of the FFT by implying increased endogeneity of the Big Five personality dimensions.

Another assumption is that of *assortative mating*, which presumes that parents do not mate based on similarity. If they did, it would affect the aforementioned normal distribution of phenotypes, which in turn would inflate the narrow heritability estimate. This assumption is repeatedly proven to be violated in several traits, such as intelligence (Vinkhuyzen et al., 2012). However, the effect of assortative mating appears minimal for personality traits (Luo & Klohnen, 2005). Therefore, this thesis assumes ignorable influences on the estimates.

From a CTD perspective, siblings can interact in mainly two ways: they can "cooperate" and increase similar behavior (creating positive interaction effects) or "compete" and increase different behavior (creating negative interaction effects) (Neale & Maes, 2004). Some CTD models include these types of interaction effects, yet these are not included in our analysis, in order to adhere to a less complex and therefore computationally feasible model. Sibling effects also bring into question the assumption of representativeness. Twin studies are criticized on the assumption that twins are representative of the population—most people do not grow up with a same-aged or identical sibling. This seems to have consequences for some traits, such as cognitive ability in early childhood (Webbink et al., 2008). Yet, evidence supports the representativeness of personality traits of twins (Johnson et al., 2002) and is an assumption we choose to make when generalizing the results.

Although these considerations might impose serious errors on the multivariate genetic estimates, they are not controlled for. The considerations should be taken into account when interpreting the results, providing an uncertainty that would accompany any method. Nonetheless, the CTD is a well-founded method, and the design applied in this thesis follows standard CTD assumptions. A satisfactory degree of the estimates' validity is therefore assumed.

4.1.3. Appraisal of the Factor Extraction Method

In this thesis, PCA was applied to extract factors. The other main method used for factor extraction is exploratory factor analysis (EFA). The two are used interchangeably as factor extraction methods, but the PCA was used more frequently when developing the NEO-PI-R and other Big Five measures (Costa & McCrae, 1992; Goldberg, 1990). Although they are often used interchangeably and yield similar results, Schreiber (2020) highlighted the mathematical differences between them. While the PCA seeks to extract maximum variance with the first components, EFA attempts to reproduce the correlation matrix with a limited set of factors. PCA assumes the factors, retained or not, account for all the variation in the components. EFA, on the other hand, assumes the observed variance to consist of common and unique variance, estimating factors only from the common variance.

This theoretical difference is the basis of the criticism regarding the use of PCA as a method for obtaining parameters that seek to reflect latent variables (Franić et al., 2014; Widaman, 1993). EFA is considered to more closely resemble latent variable theory. However, Velicer and Jackson (1990) compared principal component and factor analysis and showed the similarities and equal generalizability of the solutions, deeming the differences miniscule. Although EFA may have a favorable theoretical foundation, the inferences deduced from this method also have their limitations. Schreiber (2020) pointed out that the interpretation of results from factor analysis often are "confused with the discovery of an underlying structure for a set of variables. The interpretation is not warranted by the mathematics of factor analysis. There is a fundamental indeterminacy in factor analysis" (p. 2).

The use of PCA instead of EFA in this thesis might be considered a limitation. However, the empirical similarities and interchangeability of the two in the development of the FFM make the issue of factor retention method of lesser relevance in the interpretation of the results.

4.2. The Norwegian Big Five

The principal component analysis of the middle-aged Norwegian sample revealed five factors that simultaneously converged with and diverged from the original proposed five-factor structure. As described in the introduction, several solutions have been shown to emerge within and between cultures and would have been possible in the current sample. Needless to say, the applied instrument was constructed to measure *five* factors, and there are limitations attached to the NEO PI-R as input to the PCA. However, similar limitations have been present in studies that have failed to replicate five factors. Gurven et al. (2013) applied an instrument intended to measure five factors (a culturally and linguistically modified *Big Five Inventory*) when eleven factors were extracted from the exploratory factor analysis. As described in the introduction, robust two-factor structures were evident in diverse cultures (Saucier et al., 2014). When applying an instrument intended to measure the Big Five dimensions, one could even have expected a two-factor solution comprising the two α and β factors to emerge (Digman, 1997). Nonetheless, a five-factor structure was replicated in the Norwegian sample, as it has been in other studies that have applied facets instead of items (McCrae et al., 2001; Yamagata et al., 2006). It should be noted that studies that have applied facets, research overlapping samples. Extraction of five factors by parallel analysis in our sample thus strengthens a five-factor model in a Norwegian context.

The PCA of our sample was both similar to and different from the original FFM. The N, O, and C factors were to a large extent replicated in the current study. This supports the notion of universality. However, the A and E factors seemed unclear and overlapping, which calls for a discussion. In abstracting the semantic content of the factors, we see that they do not really align with the ideas of agreeableness or extraversion. The first A/E factor contained E1 warmth, E2 gregariousness, E6 positive emotion, A3 altruism, A1 trust, and A6 tender-mindedness. The five-factor extraversion seems to be more activity and energy-oriented than the more people-oriented cluster in our study, which might be labeled *connectedness*. The second A/E factor correlated positively with A4 compliance, A2 straightforwardness, A5 modesty, and negatively with E5 excitement-seeking, E3 assertiveness, and N5 impulsiveness. Agreeableness as proposed within the framework of the FFM seems less constraining and inhibiting than the current factor, which might be labeled *social cautiousness*. This nuances the notions of replicability and universality and falls in line with research showing varying expressions of the *interpersonal circumplex* (Rolland, 2002). Rolland (2002) has pointed to three reasons for this variation (p. 9): i) differing rotation methods, ii) the complex relationship between the interpersonal factors, and iii) real cross-cultural differences. Before discussing the proposed explanations, we begin with a short discussion of the translation.

4.2.1. Translation

We do not believe the discrepancies were caused by the Norwegian translation. A study of the Norwegian translation of the NEO PI-R reported factor congruences far above the thresholds of acceptability (Martinsen et al., 2011). Yet, facet congruences varied more, but only two facets lacked acceptable congruence. A6 tender-mindedness was not congruent with the American sample in two out of four translation studies and O6 values lacked congruence in one out of four. The Cronbach's alphas for these facets were also the lowest reported in the current sample. A6 tender-mindedness also has the lowest reported Cronbach's alpha that was reported by Costa and McCrae (1992), and is therefore considered a more fundamental validity issue rather than a translation issue. Congruence with the American instrument seems altogether satisfying and effects of the translation are deemed miniscule enough to ignore.

4.2.2. Rotation of the Interpersonal Circumplex

Rolland (2002) described how procrustes rotations generally replicate A and E, while varimax rotations vary across cultures. Procrustes rotations seek to rotate the factors as the original five-factor structure, while varimax rotations pose no expectations to the rotation of factors, adhering to the ability to explain variance and parsimony as the only criteria. The heterogeneous agreeableness and extraversion factors have also been reported in an American sample when applying confirmatory factor analysis (Vassend & Skrondal, 2011). The trend that Rolland (2002) has highlighted is evident in Norway as well. Translational studies have replicated the five factors by applying procrustes rotation (Martinsen et al., 2011), while our oblique rotation yielded a diverging factor structure.

Although the factors have theoretically been considered orthogonal, empirical evidence has suggested an oblique nature. An example is comparably superior fit indices for the oblique NEO PI-R models containing cross-loadings, reported by Vassend and Skrondal (2011). While non-orthogonality seems true for all of the five factors, it seems to specifically affect the A and E factors, which seem to vary in phenotypic expression across cultures. The cross-cultural variation has led to the interpersonal factors being described as a circumplex rather than independent factors (Rolland, 2002). To control for rotation method, we also performed an orthogonal varimax rotation. This yielded the same tendencies, indicating that the degree of correlation between factors was not the root of the discrepancy in this sample.

4.2.3. The Role of Culture and Language

Given their social aspects, it seems natural that the A and E factors might be more sensitive to cultural and lingual influences than are the more intrapersonal N, O, and C factors. In Rolland's (2002) view, this might be explained by true cultural differences. The question is nevertheless how to distinguish "true" cultural variation from merely lexical variation. A classic social constructivist viewpoint is that language is not necessarily a reflection of the culture, but also a creator of it. Yet another view is that the two may vary independently: A trait might be equally important in a culture with only one word for it as in a culture with 20 words for it. The distinction is impossible to make in this study, and we cannot conclude whether it is the Norwegian language, culture, or both that diverge from that of the United States' culture and language.

Bilingual personality research could shed light on this distinction. To our knowledge, bilingual research so far has focused on *levels* of the established five-factor measures in the same persons across languages (see Chen and Bond (2010) and McCrae et al. (1998)). We have not come across studies of personality *structure,* which our study investigates, across languages in the same persons. PCA of phenotypic matrices across languages would explain how language and culture affect the etiology of personality. To further investigate the cross-cultural aspects of personality in a Norwegian context, bottom-up psycholexical studies are also needed. Nonetheless, the social cautiousness and connectedness labeled A/E factors of the current study lead to questioning the five factor's universality.

4.2.4. Ambiguous Universality

Universality is a core argument for the interpretation of the Big Five's ontology as predominantly of endogenous origin. The divergent rotation followed a general pattern emerging in the etic literature while maintaining a sample-unique phenotypic structure.

Notably, the interpersonal circumplex was blended to such a degree that the proposed extraversion and agreeableness domains were unrecognizable. The results cast further doubt on the universality claim of the proposed FFM.

4.3. Issues With the Facets' Validity

4.3.1. The Crossing Paths of Heritability and Internal Consistency

To further elaborate on the ontology and etiology of the Big Five, we estimated the internal consistency and heritability of the five-factor facets. The universality of validity measures illuminates an important aspect of the Big Five model. Estimations are illustrated with comparisons to previous studies in Figure 10. No formal testing was applied, but from visual inspection, the estimations from our sample seem convergent with Jang et al. (1998) and Costa and McCrae (1992).

The similar degree of heritability initially solidifies the etiological and ontological notion of the facets' endogeneity. Although the estimates of facet heritability varies, both the mean estimates from our study (*M*=.41, *Ra*=.29-.56) and Jang et al.'s study (*M*=.39) resemble the estimate of .39 in Vukasović and Bratko's (2015) meta-analysis. This adds to the credibility of our results and strengthens the notion that genetics plays a key role in the formation and variance of facet dimensionality.

Cronbach's alphas of the facets in our sample varied from .46 to .84, with a median of .67, which is below the rule of thumb threshold for a unified construct (.70) (Tavakol & Dennick, 2011). A total of 21 facets (70 %) were below the threshold. The results are in line with previous research where the variation in internal consistency has formed a pattern across cultures (McCrae et al., 2011).

Estimates of internal consistency and heritability correlated moderately (r = .67). The association is correlational, and therefore directionality remains a question. Several causal relations are possible. The alphas of the facets may be caused by their heritability, i.e., highly heritable facets cause more internally consistent facets. Another possibility is the opposite: that lack of internal consistency affects heritability estimates with measurement error. The correlation might also be caused by a systematic environmental third variable, such as higher education levels, which could impose fewer environmental constraints and a greater economic capacity to enjoy genetically-inclined hobbies, thereby increasing E4 activity heritability levels while also providing a better vocabulary to coherently understand the items. The third variable might also be a latent genetic variable: A "true" genetic cluster might for example lack unity and have few heritable properties due to interaction effects at the genetic level and between genes and environment. In this instance, the association reflects the properties of the true genetic cluster and is not necessarily a third variable *problem*. Such a latent variable would be of interest in the pursuit of mapping the genome's relation to behavior but is not necessarily beneficial for a personality measure such as NEO PI-R, which is constructed to capture the stable phenotypic expression of personality.

4.3.2. The Mysterious Ways of Validity

Validity is not only a question of heritability and internal consistency, but also a question of how the facets relate to the model as a whole. Upon visual inspection, there does not seem to exist a pattern between the alphas and heritability estimates, and factor loadings. Facets with low heritability and alpha did not seem to consistently have high cross-loadings or low loadings to factors across the phenotypic, genetic, and environmental matrices. An example is A2 straightforwardness, which displayed heritability (.32) and alpha (.57) in the lower range of the distribution. Simultaneously, the A2 facet's loadings are quite clear: phenotypically, it loaded .59 on social cautiousness. In the additive genetic matrix it loaded .72 on the agreeableness dominated factor and .34 on conscientiousness (the only secondary loading above .30 across all levels of analysis). In the environmental matrix, it loaded .49 on agreeableness.

An opposite example is N5 impulsiveness, which generally did not have high loadings (maximum of .45) and loaded on four and three factors above .30 on the genetic and phenotypic levels, respectively. It serves well as an opposite example, because despite several cross-loadings, Cronbach's alpha (.62) and heritability (.42) were close to the median, which indicates a relatively internally-valid construct that lacks external validity by not fitting to the model. The N5 impulsiveness problem is well known. Zuckerman et al. (1993) reported that N5 impulsiveness loaded highest on a factor consisting mainly of conscientiousness facets and constructs.

4.3.3. Implications of the Dubious Facets

The dubious facets are certainly a limitation of the study and impose restrictions on the etiological pathways models. To investigate the question of the Big Five's ontology and etiology, we nonetheless chose to use the facets as the basis for our models. The reason is twofold: i) The use of items would require much heavier computational power. The decomposition of the phenotypic matrix through data-driven iterations with facets took several days. An identical procedure with the 240 items would require exponentially more time. We tested a model consisting of the 240 items, which only led to system crashes, indicating heavier computational power requirements. ii) Due to similar issues, previous studies have also opted for the use of facets. However, more powerful computers are available today. Therefore, our study might be somewhat more reliable than previous studies. Where our study obtained genetic and environmental correlation matrices in one procedure, Yamagata et al. (2006) had to rely on Cholesky decomposition of a 12x12 covariance matrix for each domain, computing covariation among facets across domains using Cholesky decomposition on a 24x24 covariance matrix for every possible combination. Worth noting is that the probability of iterative procedures estimating erroneous clusters increases with the amount of procedures. The use of *one* procedure therefore strengthens the reliability of the matrices, thereby improving Yamagata et al.'s methods and consolidating their results.

Thus, the use of facets may limit the validity of the etiological pathways models and our aims of assessing the fundamental interpretation of the Big Five. The NEO PI-R items or other input might have yielded a different local phenotypic structure than the solution generated by the facets, which in turn would have affected genetic and environmental structures. For example, Franić et al. (2014) replicated the FFM across all levels of analysis by applying items. By applying seemingly unreliable facets, the likely increased measurement error may have affected the search for possibly unified genetic and environmental factors. This creates uncertainty regarding the estimated etiological underpinnings. The questionable validity of the supposed building blocks might have biased the results of CPM vs. IPM analysis in favour of IPM. It may be that the lack of mediation by the factors was a function of the facets' validity. More well-defined constructs might have increased the factors' ability to explain variance. In this regard, the fundamental interpretation of the Big Five might be skewed towards the superiority of IPMs.

4.3.4. Sturdy Building Blocks?

When considering the evidence, the notion of the facets as building blocks for the domains is challenged. To a large degree, the facets are not organized in the way that is posited within the FFT. Reality seems more disarranged than theory. Even though a certain proportion of the facets meets the criteria for a unified construct, the systematic deviations from the proposed theoretical structure in the three matrices indicate a *lack* of unity in the Big Five. This questions the ontological and etiological assumption of the proposed hierarchical and all-encompassing personality structure and leads to a notion of the facets as descriptors of their own, sometimes independent of the domains. This might subsequently call for a reassessment of the facets as building blocks or at the very least a rearrangement of the taxonomy.

4.4. Etiological Exploration

Initially, we asked which environmental and genetic factors contribute to the emergence of five phenotypic personality domains. Differing from the proposed FFM, the local etiological IPM (Figure 11) represents an alternative dimensionality explaining the observed variance. The local etiological model fits the data the best, likely because it incorporates the systematic genetic and environmental influences to a greater degree. With previous studies in mind, several solutions could have emerged, but the results indicated a somewhat close overlap with the phenotypic factors. The following sections are focused on these factors and alternatives for their contextualization and explanation. However, before diving into interpretation of the results, a cautionary note must be made. Ontological inferences about the veridicality of the estimated genetic and environmental components must be made with caution, as they strictly speaking are reduced variance of quite coarse CTD estimations of a restricted set of input variables.

4.4.1. Endogenous Influences

The factors extracted from the additive genetic correlation matrix illuminate the coarse-grained genetic etiology of the Big Five dimensions. When comparing the two, the PCA of the A matrix tells a similar but somewhat different story than that of the phenotypic personality structure. The emergence of five factors replicated the results from Franić et al. (2014) and Yamagata et al. (2006) and strengthens the notion that additive genetics play a structuring role in the development of personality. The N, O, and C factors were distinctly replicated on the genetic level, with even stronger loadings between the facets and the factors. However, it is worth noting that the phenotypic correlations likely contain error, which would affect factor loadings. Nonetheless, there were evident divergences from the FFM, which either points towards a psychometric deficiency or flaws in the realist interpretation. As in the phenotypic matrix, the content of the A and E domains diverged quite considerably from the original FFM. In addition, some facets, such as O4 action and the precarious N5 impulsiveness, loaded on multiple domains. This calls for a discussion of either the psychometric model or how we interpret the Big Five.

Despite the lingering discussion, the question of which systematic processes lead to the additive genetic factor structures remains unanswered. An obvious explanation is the input of the analysis. As mentioned, the NEO PI-R facets are designed to systematically constitute the five factors. Arriving at five factors when reducing the variance in the genetic correlation matrix might therefore not be a surprise. A second possible answer is emotions and their neurology as described by affective neuroscience. By scrutinizing the behavior and neurology of rats, Panksepp (1998) has mapped and categorized the neurological underpinnings of basic mammalian emotions. Davis and Panksepp (2018) explored these neural networks in the context of the Big Five dimensions and hypothesized that each of the five domains corresponds to one or more emotional networks. Several of the networks of emotion overlapped with only *one* factor, such as the *seeking* network with openness to experience. However, all three *rage*, *sadness*, and *fear* networks corresponded to neuroticism. If the aim of personality taxonomies is to capture the endogenous structures, as proposed in the FFT, there is still a gap between the more finely-grained proposed neurobiological underpinnings and the phenotypically-derived clusters.

Biologically-focused theories of personality provide explanations of the structure of personality, but should also consider the degree of endogeneity. The impression of the genotype's dominating effect on phenotypic development stands in stark contrast to results from other methodological branches of genetic modeling. Single nucleotide polymorphisms (SNPs) are the most finely-grained measure of genetic variation widely available for behavioral research today. The SNPs represent variations in a single nucleotide and can be correlated with phenotypic outcome through genomic-relatedness-matrix restricted maximum likelihood (GREML), which is implemented in the genome-wide complex trait analysis (GCTA) software. GCTA software enables estimation of total variance in a complex trait explained by common SNP variation. A GREML-GCTA study of neuroticism by Realo et al. (2017) estimated the additive SNP-based heritability to be 15-16 %. The striking difference from the 39 % narrow-sense heritability estimate obtained in the meta-analysis of twin studies by Vukasović & Bratko (2015) is often referred to as *missing heritability*. Realo et al. (2017) suggested that this difference may be caused by i) nonadditive genetic variance, ii) the existence of rare and highly influential variants not represented in the genotype arrays, iii) epigenetic influences, or iv) the possibility that heritability estimates from twin studies are biased upwards for methodological reasons. We propose a fifth possible explanation. The missing heritability could be explained by the lack of precision in the psychometric constructs. The psychometric constructs, like the Big Five, are derived from the phenotype. Genetically-informed constructs derived from genotypic correlations might have added further precision. A construct such as neuroticism, which more closely resembles the structure found in additive correlation matrices, may add percentages to the SNP-based heritability by reducing the error of a less genetically-precise phenotypic construct. Taking FFT into account, the modeling of phenotypic factors should more closely resemble the genotypic factor solution.

4.4.2. Exogenous Influences

Vukasović & Bratko's (2015) meta-analysis estimated non-genetic contributions to measured personality to be 61 %. Despite the postulates in the FFT, there is little doubt that the environment has substantial effects on personality expression, including the Big Five. Our results are no exception, with non-shared environment estimates ranging from .44 to .71. Worth noting is that the E variance contains both systematic and unsystematic measurement error that might inflate the estimates (Keller & Coventry, 2005). We have not corrected this error. Briley and Tucker-Drob (2014) corrected environmental contributions for measurement error in their meta-analysis. The difference between the uncorrected and the corrected environmental contributions was .22. However, the method for correction is quite strict, and Briley and Tucker-Drob (2014) commented that "corrected estimates should be interpreted as lower bounds (...) and uncorrected estimates should be interpreted as upper bounds for environmentality" (p. 1309). We could therefore expect our estimates to be in the upper bounds and that the "true" estimates of environmental contributions might be somewhat lower, albeit still substantial.

Another correction to take into consideration is one proposed by McCrae et al. (2001), who corrected for systematic error of implicit personality theory. Although their choice to correct for the five factors (as the implicit personality theory) is interesting, the rationale is not compelling enough to abort examining the E matrix as it is.

Although the proportion of systematic variance explained by the E matrix factors is not as high as on the two other correlation matrices, the share is quite significant, especially when taking into consideration the relatively larger environment contributions and the fact that they likely contain measurement error.

The parallel analysis of the non-shared environmental correlation matrix yields a four factor solution, replicating the four factors above the retainment criteria in Yamagata et al.'s (2006) German and Japanese sample, but not the five in the Candian sample. Yamagata et al. selected five factors for the sake of parsimony and published the results of a five-factor rotation, whereas we stuck to the selection criteria. Yet, their data of a four-factor rotation was made available upon request. Interestingly, the E matrix of the German, Norwegian, and Japanese samples are quite distinct, but also share some features: i) All of them replicate distinct A and C factors. ii) All of them fail to crystalize E and O factors, creating a variant of a β factor, but in different ways. iii) The Norwegian and German N-factors are similar, while the Japanese lack two of the facets.

In line with Digman's (1997) proposal of a biologically-based β factor, several biological and evolutionary mechanisms have been proposed to explain the covariation between extraversion and openness to experience (see Aluja et al., 2003; Peterson et al., 2002; Schaller & Murray, 2008). However, explanations for the possible shared environmental basis seem lacking. Cross-cultural studies represent a possibility to research similarities and differences in environmental factor structures. Which features create cultural variation in environmental influences on personality?

Two main explanations are offered: i) degree of socioeconomic differentiation and ii) cultural and linguistic distinctness of cultures. With regard to the first explanation, Smaldino et al. (2019) found evidence supporting what they call the *niche diversity hypothesis*. This hypothesis states that "(...) greater diversity of social and ecological niches elicits a broader range of multivariate behavioural profiles and, hence, lower trait covariance in a population" (Smaldino et al., 2019, p. 1276). Following the reasoning of Smaldino et al., the increased variation in observed phenotypic personality structures is due to increased social and economical differentiation associated with industrialization. What would then appear as a lack of environmental crystalization in the German, Japanese, and Norwegian samples would be explained by a lesser degree of industrialization. The obvious problem with this explanation is that Germany, Japan, and Norway are not considered to be less industrialized than Canada. The alternative explanation is that different environmental factor structures are a function of cultural and linguistic variation. As discussed in the section on universality, culture and language seem to affect the psycholexically-elicited pattern of personality. Following this reasoning, cultural and lexical similarity yield more similar environmental factor structures. This might be why the Norwegian and German environmental factors have a higher degree of similarity to each other than they do to the Japanese. In addition, the fact that the NEO PI-R and the FFM are developed in an English-speaking North-American culture may be the reason that more environmental factors emerge in the Canadian sample.

4.4.3. The Sources of Personality Structure

To summarize, five partially-independent genetic sources were estimated, which point to the genetic structuring of the Big Five's phenotypic expression. However, the notion of predominantly endogenous etiology of personality traits is not supported. Not only is the environmental contribution to personality variation estimated to be massive in several studies, but is recurrently systematic and resembles the phenotypic structure, which indicates a significant exogenous structuring of personality. In this thesis as in previous studies, four environmental sources of variation were estimated, which probably points to differences in culture and language—an ontological aspect of the factors that is not described in the FFT.

4.5. Why do the Results Contradict the Realist Interpretation?

Both genetic and environmental factors affect the expression of personality. We estimated five genetic and four environmental sources of variability, which might represent an array of endogenous and exogenous influences. According to proponents of the realist interpretation, these sources are mediated by the five phenotypic factors, which makes them veridical entities. However, the results do not support such a notion. We propose two explanations for the lack of mediation: a methodological explanation and a fundamental explanation that questions the interpretation of the evidence derived from the psycholexical approach.

4.5.1. The Pathways Models in Context

This thesis has yielded further evidence unsupportive of the realist interpretation within the pathways models paradigm. The results thereby fall in line with the previous conclusions of Franić et al. (2014) and Jang et al. (2002). However, by applying the methodology of this paradigm, Johnson and Krueger (2004) and Lewis and Bates (2014) remained inconclusive about the ontology of the personality domains. The application of the methodological paradigm has several drawbacks, which in part are resolved in our study and in part remain problematic. The following section contextualizes the similarities and differences in results and conclusions within the paradigm.

First, our results support and augment the results of Franić et al. (2014) in three ways: i) Where Franić et al. tested the simple structure model (i.e., original NEOAC structure without cross-loadings) and the original model containing cross-loadings, our five-factor model was more sensitive to the local phenotypic data. This makes the test more general, because the lacking mediating effect is not due to misfit between the proposed five-factor structure and the local data. ii) Where Franić et al. applied the NEO-FFI (a reduced 60-item version of the NEO PI-R), we applied the full version (NEO PI-R) with facets instead of items as data. iii) Franić et al. used TLI and RMSEA of maximum likelihood estimations as criteria for best fit. The use of information criteria could potentially have yielded different results. The fact that it did not support the notion put forward by Franić et al. about the five factors not representing mediating factors between environmental and genetic sources of personality expression.

Second, Jang et al.'s (2002) study also investigated the unity of each domain by applying CPM vs. IPM comparisons. The advantage of Jang et al.'s analysis was the use of several IPMs that constrained sources of variance to one to three additive genetic and one to three unique environmental sources. The IPMs that model two sources of unique environmental variance and two sources of additive genetic variance fit best for every domain. This level of nuance was not present in our analysis. However, combining modeling of cross-loadings across domains and different sources of variability within each domain could potentially yield new insight to the ontology of Big Five.

Third, Johnson and Krueger's (2004) study yielded mixed results, making the authors hesitant to conclude in favor of either the realist or constructivist interpretation. The discrepancy with the current results may have methodological explanations: i) Johnson and Krueger assessed each domain with CPM vs. IPM vs. Cholesky models. Our analysis did not allow these nuances, due to the test of the complete model. Therefore, we can not conclude whether or not mediation is present by a specific underlying personality construct. It may, for example, be that extraversion and not agreeableness (as is the case in Johnson and Krueger's results) is a mediating mechanism in the interplay between genes, environment, and the phenotype. A weakness of Johnson and Krueger's design was, however, that it failed to take into account the complexity that is present between the factors in a full model and is therefore not in accordance with the massive amount of cross-loadings that empirically is present in the FFM. ii) It is worth noting that the Big Five model applied by Krueger and Johnson was constructed post hoc with varying amounts of items in each domain. Their applied model is possibly different enough from the NEO PI-R to yield diverging results. iii) It should also be noted that when using AIC, as our study did, the Cholesky models fit all domains better, indicating that no mediation described the data better than any mediation. The same was true for the entire model in our study and therefore falls in line with the general unsupportive evidence for the realist interpretation. However, the use of BIC, which favors parsimony to a higher degree than does AIC, could potentially yield different results in our study. CPMs fit E and N best when BIC was applied in Johnson and Krueger's study.

Fourth, Lewis and Bates (2014) compared several CPMs to IPMs, reporting mixed results. Although IPMs fit conscientiousness and openness to experience better, they concluded that genetic covariation within every HEXACO domain was represented by a single common genetic factor. The authors discussed whether the divergence from similar studies that apply five-factor instruments is due to the increased validity of the HEXACO model, an explanation that is relevant to the current findings' divergence from that of Lewis and Bates. The process of construing the HEXACO model was more in accordance with a bottom-up empirical process than was the NEO-PI-R. Other methodological differences might also explain the differences. i) Similar to Johnson and Krueger (2004), Lewis and Bates (2014) assessed each domain individually. The individual assessment of the domains does not easily compare to the current results. Future modeling of the full HEXACO model would therefore be of interest to researchers who seek elaborate understanding of personality with comparisons of the FFM and HEXACO. ii) Lewis and Bates applied several CPMs with different degrees of latent factor mediation. The various CPMs that fit data best, were not identical. Three of the best-fit CPMs mediated only genetic contributions and left environmental contributions to influence the facets directly. Only one of the best-fit CPMs mediated both genetic and environmental components, as in the current study.

Although there are methodological discrepancies that make our study and the four described studies difficult to compare, a certain tendency emerges. Three of the five studies indicated that IPMs fit the data better, while the two remaining implied that portions of the five-factor IPMs fit the data better. Although this seems to be in disfavor of a realist interpretation, implying the superiority of the constructivist interpretation, there are further limitations to consider. i) CTDs do not identify genes. Neither the A factors nor the *biological basis* postulated by FFM proponents are necessarily adequate representations of genetic clusters. ii) The potentially undifferentiated variance in the A and E components may conceal more nuanced structures. This would probably fit a CPM worse. A simulation study by Franić et al. (2013) showed that the conditions for the CPM became progressively worse as A, C, and E structures become more complex. iii) Refuting the CPM on the basis of an inferior fit compared to the IPM is a strict criterion. As the simulations study showed, the odds are stacked against the CPM. The complexity of the CPM and IPM is not on the level of the

Cholesky models, making neither the CPM nor the IPM an accurate enough basis on which to conclude.

Lastly, a major question is whether CPM vs. IPM comparisons truly answer the question of what the Big Five's ontology is. As pointed out, the comparisons are strict tests of the realist interpretation. In itself, a rejection of the realist notion based on these comparisons cannot be taken as a verification of the constructivist interpretation either. It seems out of touch to claim that human personality is dissociated from biological mechanisms; therefore, the question is *what* and *how* these mechanisms work. This is reflected in the diversity of personality models and their rather similar heritability estimates. However, statistical analysis does not so far support the notion that the five factors are causally efficient entities that produce behavior.

4.5.2. Interpretation of the Psycholexical Taxonomy

The process underlying the psycholexical approach might serve as an alternative explanation for the lack of mediation of environmental and genetic sources of variance by the Big Five. The psycholexical hypothesis states that the most important variations in human tendencies will be expressed through language. While the hypothesis is probably impossible to validate or falsify, it can be scrutinized. In his exhaustive critique of the five-factor approach, Block (1995) pointed out that the descriptions used in the NEO PI-R are far from context-independent. A criterion for all the descriptions is that they are understandable to American undergraduate students and might thereby lack the complexity and specificity that is needed for a true or adequate description of human personality. The simplicity is further increased by the use of *single words* that fail to embrace this complexity. The simplification might limit the potential of the lexical approach to identify possible dynamic mechanisms that human behavior, feelings, and thoughts probably emerge from.

Another aspect of the psycholexical approach that needs to be highlighted is the use of factor analysis and the interpretation of the results. Quantitative psychological research relies heavily on structural equation modeling, which assumes that latent variables affect the manifest measured variables (Bollen, 1989). The interpretation of factors that emerge through factor analysis is subject to the same assumption: that they are latent factors, which implies they play a causal role in emotion, cognition, and behavior. The assumption is widely accepted in psychological research. In many cases, the assumption might be justified, and that is when the latent variable is considered a construct—not a biological causal object. This is because what essentially is known about the relationship between the manifest and latent variables is their correlations. This implies that factor analysis is not comparable to a metal detector that identifies hidden objects; it is a statistical procedure that reduces the complexity of data. Several critics (e.g., Block (1995) and Franić et al. (2014)) have applied variations of the following example to illustrate this: Factor analysis of words that are used in commercials for electrical vehicles would yield dimensions by which cars vary: range, top speed, and acceleration. To infer anything about the internal structure and mechanics of vehicles from these factors would be beyond all reason. However, supported by evidence of heritability, universality, stability, and internal consistency, the five factors are assigned such structural-like properties by proponents of a realist interpretation. Heritability, universality, stability, and internal consistency are valuable arguments for any descriptive personality model, but this does not erase the fact that the factors are inherently correlational. The leap from clusters in the vocabulary of American undergraduates to the "biologically-based human tendencies" seems too long, and is probably important in explaining the lack of statistical mediation by the Big Five.

The weakness of the realist interpretation of the psycholexical taxonomy is also illustrated by the failure to precisely replicate the five dimensions with principal component analysis. Previously, we noted the interpersonal circumplex, and how it seems impossible to distinguish cultural from linguistic effects on the rotated solution. When agreeableness and extraversion cannot be consistently separated by statistical procedures, it seems difficult to argue that they comprise strictly separated psychological phenomena with distinct biological basis, especially when the separation cannot be done in the additive genetic matrix either. The lack of replicability might also affect the unity of these constructs, making them prone to fail the current test of mediation. The interpersonal circumplex could possibly be reconceptualized in the FFM to adhere to previously reported solutions (Rolland, 2002; Vassend & Skrondal, 2011). The A and E factors were reconceptualized to social cautiousness and connectedness in our thesis; yet, they did not mediate the environmental and genetic contributions. This does not mean that such phenomena are not—or cannot be—psychobiological. But it seems that they do not reflect mechanisms accurately described by the A and E constructs. The same challenge for extraversion is evident in the environmental matrix, where it is inseparable from openness to experience. Supposedly distinct psychological phenomena *causing* behavior should not display such dominating cross-loadings at every level of analysis. This reasoning contradicts the supposed link between linguistic clusters and the biological foundations for variation in human cognition, emotion and behavior.

4.5.3. The Ontological Status of the Big Five

Even though the estimated systematic genetic and environmental components are quite similar to the five domains, the mediating properties of the domains were not evident. In line with several previous studies of the Big Five's ontology, our CPM versus IPM comparisons undermine the realist notion of the five factors as veridical. The evidence suggests that the five dimensions remain a selection of effectively quantified outlines of how we commonly describe patterns of human behavior, thoughts, and emotions—in other words, an efficient, immediate, and coarse approximation of observed personality. But the realist notion of the five factors as causal constructs in complete etiological models of cognitive, emotional, and behavioral patterns remains unsupported. There are several considerations to be made with regard to the pathways paradigm. Its results highlight the realist proposition's reliance on the uncertain inference that information about causal entities can be extracted from the aggregation of common one-word personality descriptions.

5. Concluding Remarks

The three aims of our research illuminate important aspects of the Big Five personality traits. The results regarding the ambiguous universality and equivocal facet validity augments the indeterminacy of the Big Five taxonomy. The genetic and environmental factors reinforce this notion by indicating both endogenous and exogenous structuring of the traits. These results contribute etiological and ontological knowledge to the research and aim to refine the psychometrics of personality and the theories accompanying them. With regard to the fundamental interpretation of the dominant personality taxonomy, the current ontological status of the Big Five is limited to personality description and does not extend to causal explanations of personality.

6. References

Allport, G. W. (1931). What is a trait of personality? *The Journal of Abnormal and Social Psychology*, *25*(4), 368–372. https://doi.org/10.1037/h0075406

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*(1), 1–171. https://doi.org/10.1037/h0093360

Aluja, A., García, Ó., & García, L. F. (2003). Relationships among extraversion, openness to experience, and sensation seeking. *Personality and Individual Differences*, *35*(3), 671–680. https://doi.org/10.1016/S0191-8869(02)00244-1

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*(2), 356–366. https://doi.org/10.1037/0022-3514.86.2.356

Bateson, W., & Mendel, G. (2013). *Mendel's principles of heredity*. Dover Publications. (Original work published 1913)

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*(2), 187–215. https://doi.org/10.1037/0033-2909.117.2.187

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.

Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2002). Similarity of childhood experiences and personality resemblance in monozygotic and dizygotic twins: A test of the equal environments assumption. *Personality and Individual Differences*, *33*(2), 261–269. https://doi.org/10.1016/S0191-8869(01)00150-7

Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*(3), 243–273. https://doi.org/10.1023/A:1012294324713

Boyle, G. J. (2008). Critique of the five-factor model of personality. In G. J. Boyle & D. H.

Saklofske (Eds.), *The SAGE handbook of personality theory and assessment—Volume 1: Personality theories and models* (pp. 295–312). SAGE Publications Ltd.

Briley, D. A., & Tucker-Drob, E. M. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin*, *140*(5), 1303–1331. https://doi.org/10.1037/a0037091

Buchanan, R. (1994). The development of Minnesota Multiphasic Personality Inventory. *Journal of the History of the Behavioral Sciences*, *30*, 148–161. https://doi.org/10.1002/1520-6696(199404)30:23.0.CO;2-9

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A., & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, *297*(5582), 851–854. https://doi.org/10.1126/science.1072290

Cattell, R. B. (1944). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, *38*(4), 476. https://doi.org/10.1037/h0054116

Cattell, R. B. (1966). Validation and intensification of the sixteen personality factor questionnaire. In R. D. Savage (Ed.), *Readings in clinical psychology* (Vol. 12, pp. 241–254). Pergamon Press Ltd.

Chen, S. X., & Bond, M. H. (2010). Two languages, two personalities? Examining language effects on the expression of personality in a bilingual context. *Personality and Social Psychology Bulletin*, *36*(11), 1514–1528. https://doi.org/10.1177/0146167210385360

Costa, P. T., & McCrae, R. R. (1992). NEO PI-R professional manual. *Psychological Assessment Resources*, *396*.

Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*, 21–50. https://doi.org/10.1207/s15327752jpa6401_2

Costa, P. T., & McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment—Volume 2: Personality measurement and testing* (pp. 179–198). SAGE Publications Ltd. https://doi.org/10.4135/9781849200479.n9

Davis, K. L., & Panksepp, J. (2018). *The emotional foundations of personality: A neurobiological and evolutionary approach*. W.W Norton & Company.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*(6), 1246–1256. https://doi.org/10.1037/0022-3514.73.6.1246

Eysenck, H. J. (1963). Biological basis of personality. *Nature*, *199*(4898), 1031–1034.

Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, *13*(6), 667–673. https://doi.org/10.1016/0191-8869(92)90237-J

Eysenck, H. J., & Eysenck, S. B. G. (1991). *Manual of the Eysenck personality scale (adults)*. Hodder and Stoughton. https://doi.org/10.1037/t05461-000

Felson, J. (2014). What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption. *Social Science Research*, *43*, 184–199. https://doi.org/10.1016/j.ssresearch.2013.10.004

Fetvadjiev, V. H., & van de Vijver, F. J. R. (2015). Five factor model of personality, universality of. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 249–253). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.25067-8

Fisher, R. A. (1919). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433.

https://doi.org/10.1017/S0080456800012163

Franić, S., Borsboom, D., Dolan, C. V., & Boomsma, D. I. (2014). The Big Five personality traits: Psychological entities or statistical constructs? *Behavior Genetics*, *44*(6), 591–604. https://doi.org/10.1007/s10519-013-9625-7

Franić, S., Dolan, C. V., Borsboom, D., Hudziak, J. J., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2013). Can genetics help psychometrics? Improving dimensionality assessment through genetic factor modeling. *Psychological Methods*, *18*(3), 406–433. https://doi.org/10.1037/a0032755

Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34. https://doi.org/10.1037/0003-066X.48.1.26

Gurven, M., von Rueden, C., Massenkoff, M., Kaplan, H., & Vie, M. L. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager–farmers in the Bolivian Amazon. *Journal of Personality and Social Psychology*, *104*(2), 354–370. https://doi.org/10.1037/a0030841

Harris, J. R., Magnus, P., & Tambs, K. (2006). The Norwegian Institute of Public Health Twin Program of Research: An update. *Twin Research and Human Genetics*, *9*(6), 858–864. https://doi.org/10.1375/twin.9.6.858

Hathaway, S. R., & McKinley, J. C. (1951). *Minnesota Multiphasic Personality Inventory; manual (Revised)*. Psychological Corporation.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191–205. https://doi.org/10.1177/1094428104263675

Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T. M., Nordentoft, M., & Glenthøj, B. (2018). Heritability of schizophrenia and schizophrenia spectrum based on the Nationwide Danish Twin Register. *Biological Psychiatry*, *83*(6), 492–498. https://doi.org/10.1016/j.biopsych.2017.08.017

Jang, K. L., Livesley, W. J., Ando, J., Yamagata, S., Suzuki, A., Angleitner, A., Ostendorf, F., Riemann, R., & Spinath, F. (2006). Behavioral genetics of the higher-order factors of the Big Five. *Personality and Individual Differences*, *41*(2), 261–272. https://doi.org/10.1016/j.paid.2005.11.033

Jang, K. L., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Personality and Individual Differences*, *33*(1), 83–101.

Jang, K. L., Livesley, W. J., & Vemon, P. A. (1996). Heritability of the Big Five personality dimensions and their facets: A twin study. *Journal of Personality*, *64*(3), 577–592. https://doi.org/10.1111/j.1467-6494.1996.tb00522.x

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of personality. *Journal of Personality and Social Psychology*, *74*(6), 1556–1565. https://doi.org/10.1037/0022-3514.74.6.1556

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality*, *2*(3), 171–203. https://doi.org/10.1002/per.2410020302

John, O. P., & Robins, R. W. (1994). Traits and types, dynamics and development: No doors should be closed in the study of personality. *Psychological Inquiry*, *5*(2), 137–142. https://doi.org/10.1207/s15327965pli0502_10

Johnson, W., & Krueger, R. F. (2004). Genetic and environmental structure of adjectives describing the domains of the Big Five model of personality: A nationwide US twin study. *Journal of Research in Personality*, *38*(5), 448–472. https://doi.org/10.1016/j.jrp.2003.11.001

Johnson, W., Krueger, R. F., Bouchard, T. J., & McGue, M. (2002). The personalities of twins: Just ordinary folks. *Twin Research and Human Genetics*, *5*(2), 125–131. https://doi.org/10.1375/twin.5.2.125

Joseph, J. (1998). The equal environment assumption of the classical twin method: A critical analysis. *The Journal of Mind and Behavior*, *19*(3), 325–358. https://www.jstor.org/stable/43854406

Kandler, C., Richter, J., & Zapko-Willmes, A. (2019). The nature and nurture of HEXACO personality trait differences. *Zeitschrift Für Psychologie*, *227*(3), 195–206. https://doi.org/10.1027/2151-2604/a000378

Keller, M. C., & Coventry, W. L. (2005). Quantifying and addressing parameter indeterminacy in the classical twin design. *Twin Research and Human Genetics*, *8*(3), 201–213. https://doi.org/10.1375/twin.8.3.201

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behavior Genetics*, *23*(1), 21–27. https://doi.org/10.1007/BF01067551

Lewis, G. J., & Bates, T. C. (2014). How genes influence personality: Evidence from multi-facet twin analyses of the HEXACO dimensions. *Journal of Research in Personality*, *51*, 9–17. https://doi.org/10.1016/j.jrp.2014.04.004

Loehlin, J. C., & Martin, N. G. (2001). Age changes in personality traits and their heritabilities during the adult years: Evidence from Australian twin registry samples. *Personality and Individual Differences*, *30*(7), 1147–1160.

https://doi.org/10.1016/S0191-8869(00)00099-4

Loehlin, J. C., McCrae, R. R., Costa, P. T., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, *4*(32), 431–453.

Luo, S., & Klohnen, E. C. (2005). Assortative mating and marital quality in newlyweds: A couple-centered approach. *Journal of Personality and Social Psychology*, *88*(2), 304–326. https://doi.org/10.1037/0022-3514.88.2.304

Magnus, P., Berg, K., & Nance, W. E. (1983). Predicting zygosity in Norwegian twin pairs born 1915–1960. *Clinical Genetics*, *24*(2), 103–112. https://doi.org/10.1111/j.1399-0004.1983.tb02220.x

Martinsen, Ø. L., Nordvik, H., & Østbø, L. E. (2011). The NEO PI-R in a North European context. *Scandinavian Journal of Organizational Psychology*, *3*(2), 58–75. http://hdl.handle.net/11250/93444

McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139–153). Guilford.

McCrae, R. R., & Costa, P. T. (2008). Empirical and theoretical status of the five-factor model of personality traits. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment—Volume 1: Personality theories and models* (pp. 273–294). SAGE Publications Ltd.

McCrae, R. R., Jang, K. L., Livesley, W. J., Riemann, R., & Angleitner, A. (2001). Sources of structure: Genetic, environmental, and artifactual influences on the covariation of personality traits. *Journal of Personality*, *69*(4), 511–535. https://doi.org/10.1111/1467-6494.694154

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*(1), 28–50. https://doi.org/10.1177/1088868310366253

McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, *88*(3), 547–561. https://doi.org/10.1037/0022-3514.88.3.547

McCrae, R. R., Yik, M. S. M., Trapnell, P. D., Bond, M. H., & Paulhus, D. L. (1998). Interpreting personality profiles across cultures: Bilingual, acculturation, and peer rating studies of Chinese undergraduates. *Journal of Personality and Social Psychology*, *74*(4), 1041–1055. https://doi.org/10.1037/0022-3514.74.4.1041

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology*, *70*(3), 552–566. https://doi.org/10.1037/0022-3514.70.3.552

Munafò, M. R., Durrant, C., Lewis, G., & Flint, J. (2009). Gene × environment interactions at the serotonin transporter locus. *Biological Psychiatry*, *65*(3), 211–219. https://doi.org/10.1016/j.biopsych.2008.06.009

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8

Neale, M. C., & Maes, H. H. M. (2004). *Methodology for genetic studies of twins and families*. Kluwer Academic Publishers.

Norman, W. T. (1964). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, *66*(6), 574–583. https://doi.org/10.1037/h0040291

Okbay, A., Baselmans, B. M. L., De Neve, J.-E., Turley, P., Nivard, M. G., Fontana, M. A., Meddens, S. F. W., Linnér, R. K., Rietveld, C. A., Derringer, J., Gratten, J., Lee, J. J., Liu, J. Z., de Vlaming, R., Ahluwalia, T. S., Buchwald, J., Cavadino, A., Frazier-Wood, A. C., Furlotte, N. A., … Cesarini, D. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, *48*(6), 624–633. https://doi.org/10.1038/ng.3552

Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. CreateSpace Independent Publishing Platform.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*(3), 524–539. https://doi.org/10.1037/0022-3514.81.3.524

Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? Plenty! *Journal of Personality*, *68*(5), 821–835. https://doi.org/10.1111/1467-6494.00117

Peterson, J. B., Smith, K. W., & Carson, S. (2002). Openness and extraversion are associated with reduced latent inhibition: Replication and commentary. *Personality and Individual Differences*, *33*(7), 1137–1147. https://doi.org/10.1016/S0191-8869(02)00004-1

Plomin, R. R. (2013). *Behavioral genetics* (6th ed.). Worth Publishers.

Realo, A., van der Most, P. J., Allik, J., Esko, T., Jeronimus, B. F., Kööts-Ausmees, L., Mõttus, R., Tropf, F. C., Snieder, H., & Ormel, J. (2017). SNP-based heritability estimates of common and specific variance in self- and informant-reported neuroticism scales. *Journal of Personality*, *85*(6), 906–919.

https://doi.org/10.1111/jopy.12297

Rijsdijk, F. (2005). Independent pathway model. In B. S. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science*. American Cancer Society. https://doi.org/10.1002/0470013192.bsa298

Robertson, A. (1959). The sampling variance of the genetic correlation coefficient. *Biometrics*, *15*(3), 469–485. https://doi.org/10.2307/2527750

Rolland, J.-P. (2002). The cross-cultural generalizability of the five-factor model of personality. In R. R. McCrae & J. Allik (Eds.), *The five-factor model of personality across cultures* (pp. 7–28). Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/978-1-4615-0763-5_2

Rolland, J.-P., Parker, W. D., & Stumpf, H. (1998). A psychometric examination of the French translations of NEO-PI-R and NEO-FFI. *Journal of Personality Assessment*, *71*(2), 269–291. https://doi.org/10.1207/s15327752jpa7102_13

Røysamb, E., Nes, R. B., Czajkowski, N. O., & Vassend, O. (2018). Genetics, personality and wellbeing. A twin study of traits, facets and life satisfaction. *Scientific Reports*, *8*(1), Article 12298. https://doi.org/10.1038/s41598-018-29881-x

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

Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, *69*(6), 847–879. https://doi.org/10.1111/1467-6494.696167

Saucier, G., Thalmayer, A. G., Payne, D. L., Carlson, R., Sanogo, L., Ole‐Kotikash, L., Church, A. T., Katigbak, M. S., Somer, O., Szarota, P., Szirmák, Z., & Zhou, X. (2014). A basic bivariate structure of personality attributes evident across nine languages. *Journal of Personality*, *82*(1), 1–14. https://doi.org/10.1111/jopy.12028

Schaller, M., & Murray, D. R. (2008). Pathogens, personality, and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology*, *95*(1), 212–221. https://doi.org/10.1037/0022-3514.95.1.212

Schmitt, D., Allik, J., McCrae, R. R., Benet, V., Alcalay, L., Ault, L., Austers, I., Bennett, K., Bianchi, G., Boholst, F., Cunen, M., Braeckman, J., Jr, E., Caral, L. G., Caron, G., Casullo, M., Cunningham, M., Daibo, I., de backer, C., & Zupanèiè, A. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, *38*, 173–212.

Schreiber, J. B. (2020). Issues and recommendations for exploratory factor analysis and principal component analysis. *Research in Social and Administrative Pharmacy*. Advance online publication. https://doi.org/10.1016/j.sapharm.2020.07.027

Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D. I., Cornes, B. K., Davis, C., Dunkel, L., Lange, M. de, Harris, J. R., Hjelmborg, J. V. B., Luciano, M., Martin, N. G., Mortensen, J., Nisticò, L., Pedersen, N. L., Skytthe, A., Spector, T. D., Stazi, M. A., Willemsen, G., & Kaprio, J. (2003). Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Research and Human Genetics*, *6*(5), 399–408. https://doi.org/10.1375/twin.6.5.399

Smaldino, P. E., Lukaszewski, A., von Rueden, C., & Gurven, M. (2019). Niche diversity can explain cross-cultural differences in personality structure. *Nature Human Behaviour*, *3*(12), 1276–1283. https://doi.org/10.1038/s41562-019-0730-3

Smith, D. J., Escott-Price, V., Davies, G., Bailey, M. E. S., Colodro-Conde, L., Ward, J., Vedernikov, A., Marioni, R., Cullen, B., Lyall, D., Hagenaars, S. P., Liewald, D. C.

M., Luciano, M., Gale, C. R., Ritchie, S. J., Hayward, C., Nicholl, B., Bulik-Sullivan, B., Adams, M., … O'Donovan, M. C. (2016). Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Molecular Psychiatry*, *21*(6), 749–757. https://doi.org/10.1038/mp.2016.49

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tellegen, A., Lykken, D. T., Bouchard, T. J., Wilcox, K. J., Segal, N. L., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology*, *54*(6), 1031–1039. https://doi.org/10.1037/0022-3514.54.6.1031

van den Berg, S. M., de Moor, M. H. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Arias Vasquez, A., Matteson, L. K., Derringer, J., Esko, T., Amin, N., Gordon, S. D., Hansell, N. K., Hart, A. B., Seppälä, I., Huffman, J. E., Konte, B., Lahti, J., Lee, M., Miller, M., … Boomsma, D. I. (2016). Meta-analysis of genome-wide association studies for extraversion: Findings from the Genetics of Personality Consortium. *Behavior Genetics*, *46*(2), 170–182. https://doi.org/10.1007/s10519-015-9735-5

Vassend, O., Røysamb, E., Nielsen, C. S., & Czajkowski, N. O. (2017). Musculoskeletal complaints, anxiety–depression symptoms, and neuroticism: A study of middle-aged twins. *Health Psychology*, *36*(8), 729.

Vassend, O., & Skrondal, A. (2011). The NEO personality inventory revised (NEO-PI-R): Exploring the measurement structure and variants of the five-factor model. *Personality and Individual Differences*, *50*(8), 1300–1304. https://doi.org/10.1016/j.paid.2011.03.002

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, *25*(1), 1–28. https://doi.org/10.1207/s15327906mbr2501_1

Vinkhuyzen, A. A. E., van der Sluis, S., Maes, H. H. M., & Posthuma, D. (2012). Reconsidering the heritability of intelligence in adulthood: Taking assortative mating and cultural transmission into account. *Behavior Genetics*, *42*(2), 187–198. https://doi.org/10.1007/s10519-011-9507-9

Vukasović, T., & Bratko, D. (2015). Heritability of personality: A meta-analysis of behavior genetic studies. *Psychological Bulletin*, *141*(4), 769–785. https://doi.org/10.1037/bul0000017

Waldherr, A., & Muck, P. M. (2011). Towards an integrative approach to communication styles: The interpersonal circumplex and the five-factor theory of personality as frames of reference. *Communications*, *36*(1), 1–27. https://doi.org/10.1515/comm.2011.001

Webbink, D., Posthuma, D., Boomsma, D. I., de Geus, E. J. C., & Visscher, P. M. (2008). Do twins have lower cognitive ability than singletons? *Intelligence*, *36*(6), 539–547. https://doi.org/10.1016/j.intell.2007.12.002

Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*(3), 263–311. https://doi.org/10.1207/s15327906mbr2803_1

Widiger, T. A. (2017). *The Oxford handbook of the five factor model*. Oxford University Press. DOI: 10.1093/oxfordhb/9780199352487.001.0001

Wright, S. (1921). Systems of mating. V. General considerations. *Genetics*, *6*(2), 167–178. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1200505/

Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics*, *5*(3), 161–215. https://www.jstor.org/stable/2957502

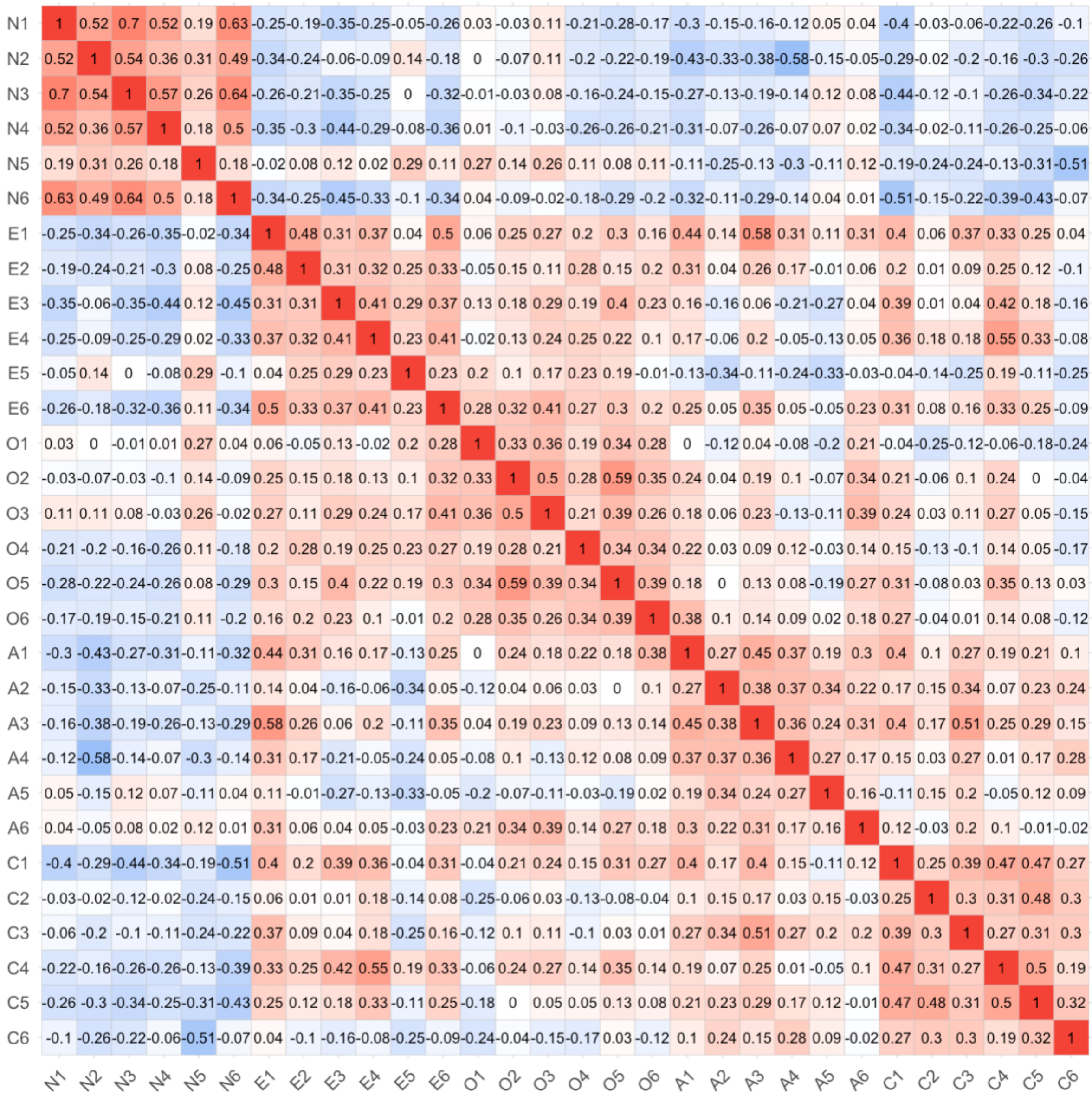
Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., Ostendorf, F., Angleitner, A., Riemann, R., Spinath, F. M., Livesley, W. J., & Jang, K. L. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*(6), 987–998. https://doi.org/10.1037/0022-3514.90.6.987

Zuckerman, M., Kuhlman, D. M., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology*, *65*(4), 757–768. https://doi.org/10.1037/0022-3514.65.4.757

Appendix A

Figure 12

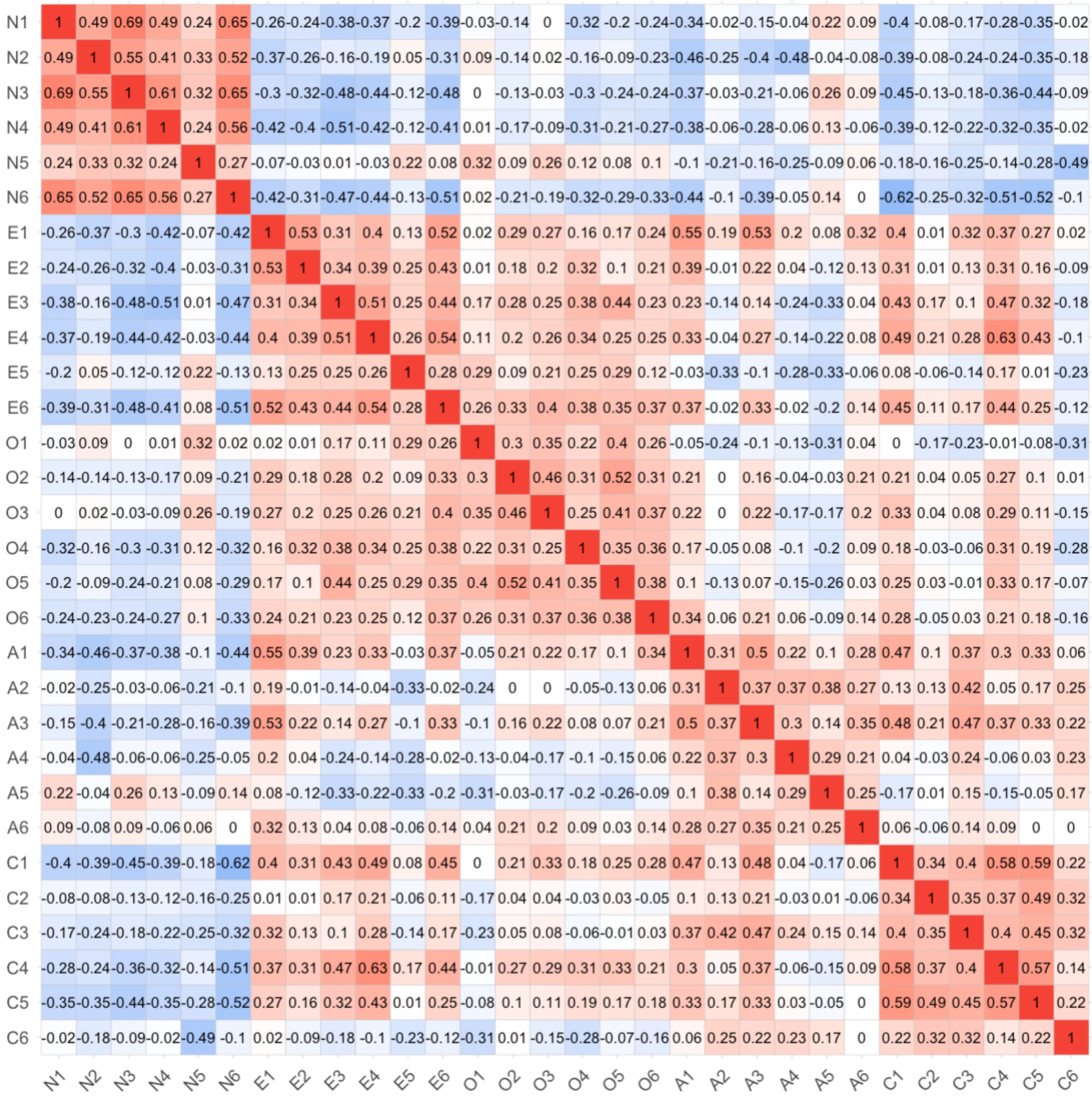
Female MZ twin correlations



Note. Female correlation matrices were used to assess inclusion of C and D components due to encompassment of a higher proportion of the participants.

Figure 13

Female DZ twin correlations



Appendix B

Figure 14

Parallel analysis scree plot of the phenotypic correlation matrix

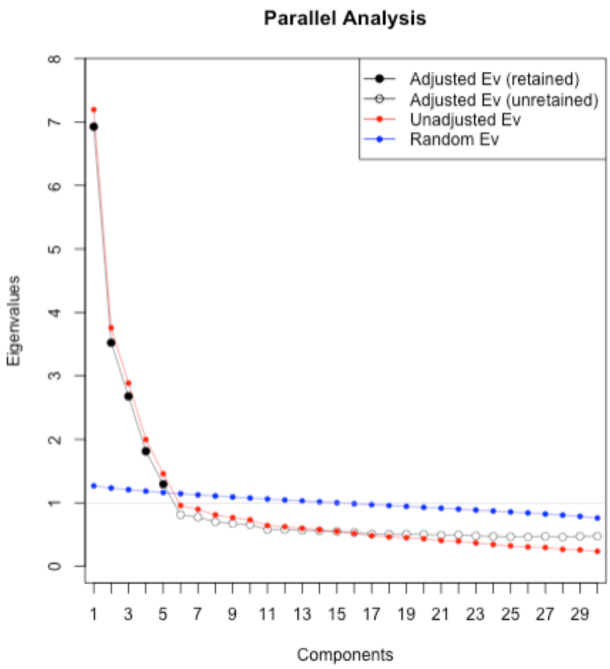


Figure 15

Parallel analysis scree plot of the additive genetic correlation matrix

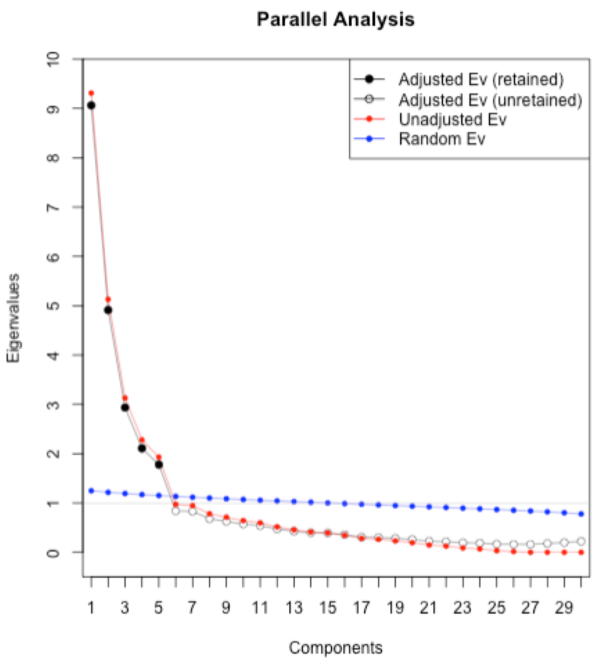


Figure 16

Parallel analysis scree plot of the nonshared environmental correlation matrix